

ART SEARCH RECORD

Patent Examiner Francis Moonan
US Patent Application: 09/622978
Inventor: Hull, Roger; Harper, Glyn

I. EAST Search.

Databases=USPAT; US-PGPUB; EPO; JPO; DERW
Date= 8 November 2001

Query=(Schenk.IN. or Sagi.IN. or Remy.IN. or Swennen.IN. or
Dietzen.IN. or Geering.IN. or McMichael.IN. or thomas.IN or Grof.IN or
Elliott.IN. or Hull.IN or Harper.IN or Lockhart.IN.)and (sugarcane or
banana or plantain)

Results= 4

Art Considered= 4

Adams et al. 23 February 1999.
Methods and compositions for the production of stablytransformed fertile
monocot plants and cells thereof.
US Patent No. 5,874,265

Olszewski et al. 25 February 1999.
Sugarcane bacilliform virus promoter.
WIPO No. WO 99/09190

Olszewski et al. 30 November 1999.
Sugarcane bacilliform virus promoter.
US Patent No. 5,994,123

Olszewski et al. 25 July 2000.
Sugarcane bacilliform virus promoter.
US Patent No. 5,994,123

II. ~~DIALOG~~/STN Search.

Databases= AGRICULTURE
Date=9 November 2001

Query=Schenk or Sagi or Remy or Swennen or Dietzen or Geering or
McMichael or Thomas or Grof or Elliott or Hull or Harper or Lockhart
and sugarcane or banana or plantain not mammal? not enzym? and virus
or viral and pararetrovirus or badnavirus

Results= 146

Art Considered= 14

L8 ANSWER 1 OF 146 CAPLUS COPYRIGHT 2001 ACS

ACCESSION NUMBER: 2001:763212 CAPLUS

TITLE: Eukaryotic recombinant bidirectional transcription
vector to effect gene silencing via dsRNA-mediated
inhibition of gene expression

INVENTOR(S): Palmer, Kenneth E.; Pogue, Gregory P.

PATENT ASSIGNEE(S): Large Scale Biology Corporation, USA

SOURCE: PCT Int. Appl., 73 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
------------	------	------	-----------------	------

WO 2001077350	A2	20011018	WO 2001-US11436	20010404
---------------	----	----------	-----------------	----------

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN,
CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM,
HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS,
LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO,
RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ,
VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM

RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY,
DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF,
BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: US 2000-545574 A 20000407

L8 ANSWER 14 OF 146 BIOSIS COPYRIGHT 2001 BIOSIS DUPLICATE 6

ACCESSION NUMBER: 2000:216163 BIOSIS

DOCUMENT NUMBER: PREV200000216163

TITLE: The rice tungro bacilliform virus gene II product interacts
with the coat protein domain of the viral gene III
polyprotein.

AUTHOR(S): Herzog, Etienne; Guerra-Peraza, Orlene; Hohn, Thomas (1)

CORPORATE SOURCE: (1) Friedrich Miescher Institute, Maulbeerstrasse 66,
CH-4058, Basel Switzerland

SOURCE: Journal of Virology, (March, 2000) Vol. 74, No. 5, pp.
2073-2083.

ISSN: 0022-538X.

DOCUMENT TYPE: Article

LANGUAGE: English

SUMMARY LANGUAGE: English

L8 ANSWER 15 OF 146 AGRICOLA DUPLICATE 7

ACCESSION NUMBER: 2001:28866 AGRICOLA

DOCUMENT NUMBER: IND22298361

TITLE: Characterization and genomic analysis of tobacco vein
clearing virus, a plant pararetrovirus that
is transmitted vertically and related to sequences
integrated in the host genome.

AUTHOR(S): Lockhart, B.E.; Menke, J.; Dahal, G.; Olszewski, N.E.

AVAILABILITY: DNAL (QR360.A1J6)
SOURCE: The Journal of general virology, June 2000. Vol. 81,
No. pt.6. p. 1579-1585
Publisher: Reading : Society for General Microbiology.
CODEN: JGVIA Y; ISSN: 0022-1317
NOTE: Includes references
PUB. COUNTRY: England; United Kingdom
DOCUMENT TYPE: Article
FILE SEGMENT: Non-U.S. Imprint other than FAO
LANGUAGE: English

L8 ANSWER 25 OF 146 AGRICOLA DUPLICATE 13
ACCESSION NUMBER: 2000:46607 AGRICOLA
DOCUMENT NUMBER: IND22046496
TITLE: Integrated pararetroviral sequences define a unique
class of dispersed repetitive DNA in plants.
AUTHOR(S): Jakowitsch, J.; Mette, M.F.; Winden, J. van der.;
Matzke, M.A.; Matzke, A.J.M.
CORPORATE SOURCE: Institute of Molecular Biology, Salzburg, Austria.
SOURCE: Proceedings of the National Academy of Sciences of the
United States of America, Nov 9, 1999. Vol. 96, No.
23. p. 13241-13246
Publisher: Washington, D.C. : National Academy of
Sciences,
CODEN: PNAS A6; ISSN: 0027-8424
NOTE: Includes references
PUB. COUNTRY: District of Columbia; United States
DOCUMENT TYPE: Article; Conference
FILE SEGMENT: U.S. Imprints not USDA, Experiment or Extension
LANGUAGE: English

L8 ANSWER 26 OF 146 CAPLUS COPYRIGHT 2001 ACS DUPLICATE 14
ACCESSION NUMBER: 1999:523005 CAPLUS
DOCUMENT NUMBER: 131:267862
TITLE: A short open reading frame terminating in front of a
stable hairpin is the conserved feature in pregenomic
RNA leaders of plant pararetroviruses
AUTHOR(S): Pooggin, Mikhail M.; Futterer, Johannes; Skryabin,
Konstantin G.; Hohn, Thomas
CORPORATE SOURCE: Friedrich Miescher Institute, Basel, CH-4002, Switz.
SOURCE: J. Gen. Virol. (1999), 80(8), 2217-2228
CODEN: JGVIA Y; ISSN: 0022-1317
PUBLISHER: Society for General Microbiology
DOCUMENT TYPE: Journal
LANGUAGE: English
REFERENCE COUNT: 54
REFERENCE(S): (1) Bacharach, E; Journal of Virology 1998, V72, P6944
CAPLUS
(2) Bao, Y; Virology 1993, V197, P445 CAPLUS
(3) Bhattacharyya-Pakrasi, M; Plant Journal 1993, V4,
P71 CAPLUS
(4) Bonneville, J; Cell 1989, V59, P1135 CAPLUS
(5) Bouhida, M; Journal of General Virology 1993, V74,
P15 CAPLUS
ALL CITATIONS AVAILABLE IN THE RE FORMAT

L8 ANSWER 27 OF 146 Elsevier BIOBASE COPYRIGHT 2001 Elsevier Science B.V.

DUPLICATE

ACCESSION NUMBER: 1999132177 Elsevier BIOBASE

TITLE: A promoter from sugarcane bacilliform
badnavirus drives transgene expression in
banana and other monocot and dicot plants

AUTHOR: Schenk P.M.; Sagi L.; Remans T.; Dietzgen R.G.;
Bernard M.J.; Graham M.W.; Manners J.M.

CORPORATE SOURCE: P.M. Schenk, Coop. Res. Trop. Plant Pathol. Ctr.,
University of Queensland, John Hines Building, St.
Lucia, QLD 4072, Australia.

SOURCE: Plant Molecular Biology, (1999), 39/6 (1221-1230), 39
reference(s)

CODEN: PMBIDB ISSN: 0167-4412

DOCUMENT TYPE: Journal; Article

COUNTRY: Netherlands

LANGUAGE: English

SUMMARY LANGUAGE: English

L8 ANSWER 29 OF 146 AGRICOLA DUPLICATE 17

ACCESSION NUMBER: 2000:15812 AGRICOLA

DOCUMENT NUMBER: IND22025459

TITLE: Structure and promoter/leader deletion analysis of
mirabilis mosaic virus (MMV) full-length transcript
promoter in transgenic plants.

AUTHOR(S): Dey, N.; Maiti, I.B.

CORPORATE SOURCE: University of Kentucky, Lexington, KY.

AVAILABILITY: DNAL (QK710.P62)

SOURCE: Plant molecular biology, July 1999. Vol. 40, No. 5. p.
771-782

Publisher: Dordrecht : Kluwer Academic Publishers.

CODEN: PMBIDB; ISSN: 0167-4412

NOTE: Includes references

PUB. COUNTRY: Netherlands

DOCUMENT TYPE: Article

FILE SEGMENT: Non-U.S. Imprint other than FAO

LANGUAGE: English

L8 ANSWER 106 OF 146 AGRICOLA DUPLICATE 57

ACCESSION NUMBER: 93:51829 AGRICOLA

DOCUMENT NUMBER: IND93031884

TITLE: An analysis of the complete sequence of a sugarcane
bacilliform virus genome infectious to banana and
rice.

AUTHOR(S): Bouhida, M.; Lockhart, B.E.L.; Olszewski, N.E.

CORPORATE SOURCE: Institut Agronomique Hassan II, Agadir, Morocco

AVAILABILITY: DNAL (QR360.A1J6)

SOURCE: The Journal of general virology, Jan 1993. Vol. 74,
No. pt.1. p. 15-22

Publisher: Reading : Society for General Microbiology.

CODEN: JGVIA Y; ISSN: 0022-1317

NOTE: Includes references.

DOCUMENT TYPE: Article

FILE SEGMENT: Non-U.S. Imprint other than FAO
LANGUAGE: English

L8 ANSWER 134 OF 146 GENBANK.RTM. COPYRIGHT 2001
LOCUS (LOC): NTA414165 GenBank (R)
GenBank ACC. NO. (GBN): AJ414165
CAS REGISTRY NO. (RN): 361143-48-8
SEQUENCE LENGTH (SQL): 4806
MOLECULE TYPE (CI): DNA; linear
DIVISION CODE (CI): Plants, fungi, algae
DATE (DATE): 3 Oct 2001
DEFINITION (DEF): Nicotiana tabacum ORF1, ORF2 and partial ORF3,
pararetrovirus-like sequence, V2.
KEYWORDS (ST): coat protein; movement protein; ORF1; ORF2; ORF3;
polyprotein
SOURCE: common tobacco.
ORGANISM (ORGN): Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Tracheophyta; Spermatophyta; Magnoliophyta;
eudicotyledons; core eudicots; Asteridae; euasterids I;
Solanales; Solanaceae; Nicotiana
NUCLEIC ACID COUNT (NA): 2445 a 542 c 699 g 1120 t
REFERENCE: 1 (bases 1 to 4806)
AUTHOR (AU): Jakowitsch,J.; Mette,M.F.; van Der Winden,J.;
Matzke,M.A.; Matzke,A.J.
TITLE (TI): Integrated pararetroviral sequences define a unique
class of dispersed repetitive DNA in plants
JOURNAL (SO): Proc. Natl. Acad. Sci. U.S.A., 96 (23), 13241-13246
(1999)
OTHER SOURCE (OS): CA 132:103712
REFERENCE: 2 (bases 1 to 4806)
AUTHOR (AU): van der Winden,J.
TITLE (TI): Direct Submission
JOURNAL (SO): Submitted (25-SEP-2001) van der Winden J., Plant
Molecular Genetics, Institute of Molecular Biology,
Billrothstrasse 11, 5020 Salzburg, AUSTRIA

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..4806	/organism="Nicotiana tabacum" /db-xref="taxon:4097"
mRNA	202..1998	/note="ORF1"
exon	202..1998	/note="ORF1"
CDS	202..1998	/number=1 /note="ORF1" /codon-start=1 /product="putative coat protein" /protein-id="CAC88783.1" /db-xref="GI:15963357" /translation="MNKEEFALEEKTYENPEGLK ITIIFSNLGRYKKIGNNLNLMLE KETVKLEDSLAMVRITKENEEIDRKREIKEIQQ QAKEKIQQIEEVKNTKITELEKEL EMLKQMYANKLKEKEKRKEEEEELKLTNEIERFK

LQLEEVQESPSKISINEINDQSDK
 DTELGENYSETSETYTELIEKTEKIKINPEINTG
 DMNEDKPSISGIKNPKQLNPTYR
 VSYDAYDRNKTLWDKRLNKKWAPRQITEQYNFLD
 LDCVADINKTIQLWIGYISKQLID
 NKITITETPGYIERTLIGTVKLWLQNLSSSLDT
 LRSNKKLDGTTTTTVTDILNKYEI
 AIRNEFSSMTTEVEEQNKEKITNRNLMTKLAICN
 MCYIDEYTCAFRDYYYKGTYSPE
 SKEIRKLYFTKLPEPFSSKIIKSWNEAGLADTLG
 VRIKFLQNWFIELCEKYKENMKME
 KILVKNLACCKSRIAPQGCTDKYYKKEGKKKKF
 KSKYSKYKYPKRRRYVKNYKHK
 KPYRKKKKLTECTCYNCGKLGHAKDCKLPKNPK
 KKQITEILIDNDKYTQVEYVDYEL
 SSEDSEMYEISENEFESEINKDIEESDEENYD"

mRNA	1979..2932	/note="ORF2"
CDS	1979..2932	/note="ORF2 truncated"
		/codon-start=1
		/product="putative movement protein"
		/protein-id="CAC88784.1"
		/db-xref="GI:15963358"
		/translation="MKKIMTEKDIQIIQQEEHQD EQSSEQKIIFDANIFEQIKGKELD LSIDKILEVPTIKNWFKRQKEEYVVSQREHIID CKYIKGKAQIPILNKRLLNKEIQD IKAKNPIKYVHLGGTEILTKACFREGIDTPIEY LADDRIVQPIEKSIIASAVRGNLIY KKLKSIVSANYSAVDDKNIDKSLVLYWKMSGIE LAPGSKIFTARCKNLYVLTTKHKI TAKNKINKIKIENPFERIVSIDNNDYSYTEIDM DEDLEIVKERLSTSKRINNEMPET SSRSTSRSTSKRINYTPQKLIEQKIEINPHHY YITGIMD"
exon	1979..2932	/note="ORF2"
		/number=1
exon	3256..>4806	/note="ORF3"
		/number=1
mRNA	3256..>4806	/note="ORF3"
CDS	3256..>4806	/note="ORF3 aspartic proteinase, reverse transcriptase, ribonuclease H"
		/codon-start=1
		/product="putative polyprotein"
		/protein-id="CAC88785.1"
		/db-xref="GI:15963359"
		/translation="MPKIYILSKIIVEGYYNRY TPMVDTGAEANMCRHNCPLPESKWE KLKNPIVVTGFNNEGSMITYKARNIKIQIWDKIL TIEEISYEFTNKDILLGMPFLDK LYPHIITKTHWWFTTPCKQKLGAKRNNKVRKTT PWIKGSEKITQKLENIIQSNHNIE IIIFSINKIKPLQDKLELLYNDNPLQGWEKHQTK IKIELIDENSIITQKPLKYNFNDL TEFKMHIKELLDNNYIQESNSKHTSPAFIVNKH

EQKRGKSRMVIDYRNLNAKTKTYN
 YPIPKNILKIRQIQGYNYFSKFDYKSGFYHLKLE
 DESKKLTAFTVPQGFYEWNVLPFG
 YKNAPGRYQHFMDNYFNQLENCIYIDDILLYSR
 TENEHIKLEKFIHIVEISGISLS
 KKKA EVMKNQIEFLGIQMDKNGIKMQTHVVQKII
 NLNETLDTKKKLQSFLGLVNQVRE
 YIPKLAENLKLPLQKKLKKDIEYHFDEKDKIHIQK
 IKNMCKKLPKLYFPDEKKQFTYIV
 ETDSSDRVD"

SEQUENCE (SEQ):

1 cattaatagt acaaacccgc aaaacagtga aacaatcaaa ataactaaga aatcgcgagaa
 61 aattaggaag aaactaaaga aactatata agaatatgaa tacttaagta tcacaaaaat
 121 aaacaaagct aggttgcca aattaatcga tataatatct aaacagaaat ataatatgt
 181 ttgtatcta aaggataaaa gatgaataaa gaagaattcg cattagaaga gaaaacatat
 241 gagaatccag aaggattaaa aataacaata atattttcta acttaggaag aagatataaa
 301 aaaataggaa ataacctaaa cttaattgta gaaaaagaaa ctgtaaaact agaggatagt
 361 ttaaccgccca tggtagaata acaaaaagaa aacgaagaaa tagatagaaa acgagagatt
 421 aaagaaatc aacagcaagc taaagaaaa atacaacaga tagaggaagt aaaaaacact
 481 aaataacag aattagaaaa agaattagag atgctaaaac agatgtatgc aaataaacta
 541 aaagaaaagg aaaaacgtaa agaagaagaa gaagaactaa aactaacaag tgagatagaa
 601 agattcaaat tacagttaga agaagtacag gagagtccat caaaaataag tataaacgaa
 661 ataatgacc aaagtataa agacacagaa ctaggagaaa actattcaga aacaagtga
 721 acatacacag aacttataga aaaaacagaa aagataaaaa taaaccaga aataatata
 781 ggagatatga acgaagataa accaagcata tcaggaataa aaatccaaa acaattaac
 841 ccaacctatt acagagtaag ttatgatgca tatgacagaa aaaaacatt atgggataaa
 901 aggttaataa agaatgggc accaagacag ataactgaac aatataatt ttatgatcta
 961 gattgtgtag cagatataaa taaacaata caattatgga taggatata ctcaaaaca
 1021 ctaatagata ataaataac aataacggaa acaccaggat atagagaa aacattaata
 1081 ggaactgtaa aattatggtt acaaaatta tcaagtgaag gcttagacac attaggagt
 1141 aataaaaaac ttgacgttac aactacaaca acagttacag atatatata taaatatgaa
 1201 atagcaataa gaaatgaatt tagtagtat acaacagaag tagaagaaca aaataaagaa
 1261 aaaattaca atagaaattt aatgacaaa tttagcaat gtaatatgt ttatatagat
 1321 gaatatatt gtgcatttag agactattat tataaaggga catatagtc agatgaaagt
 1381 aaagagataa gaaattata tttaaaaa ttaccagaac ccttagctc aaaaataata
 1441 aaaagttaga acgaagcagg actagcagat acattaggag taaggataaa atttctaca
 1501 aactggttta tagaattatg tgaaaaatat aaagaaaaca tgaattgga aaaaatatta
 1561 gtaaaaaatt tggcatgttg caaaagtaga atagcacccc aatttgctg cacagataaa
 1621 tattacaaa aagaaggaaa gaaaaagaaa tttaataca aatttcaaa gtataaatat
 1681 aggaaccaa gaagaagata ttatgtaaa aattataac ataaaaacc atacaggaaa
 1741 aagaaaaaac taacagaatg tacttgctat aattgtgga agctaggaca cttagccaaa
 1801 gattgtaaa taccaaaaaa ccaagaag aaacaaatta ccgaatatt gatagataat
 1861 gataaatata cacaagtaga atagtagat tatgaattaa gcagcgaaga cagcatgtat
 1921 gagatatcag aaaacgagtt ctctgaaagt gaaataaaca aggatataga agagtcagat
 1981 gaagaaaaat atgactgaaa aagatatata aattataca caagaagaac accaagatga
 2041 acaatctca gaacaaaaa taatattga cgctaatata ttgaacaaa taaaaggaaa
 2101 agaattgga ctaagttag acaaatatt agaagtacca acaataaaga attggttaa
 2161 aagacaaaaa gaagaatact atgtagtaag ccaagagaa catataatg attgtaata
 2221 catcaaggt aaagcacaaa taccaattt aaataaaga ctactaaata aagaatata
 2281 agatataaaa gcaaaaaatc caataaata tgtacactta ggaggacgg agatcctaac
 2341 aaaagcatgt tttagggaag gaatagatac ccctatagaa atatactgg cagatgatag
 2401 aattgtaca cctatagaa agagtataat aagtgtgta agaggtaact taatatata
 2461 aaattaaaa tctatagtaa gtgctaacta tttagtagca gtagatgata aaatataga
 2521 taaatcatta gtattatact ggaaatgtc tggaatagaa ttgaccag gaagtaaat
 2581 attcacagca agatgtaaa atctatatgt cttaacaaca aaacataaga taacagctaa

2641 aaataaaata aataaaataa aaatagaaaa tccattcgaa agaatagtat cagttataga
 2701 caacaatgat tacagttata cagaaataga catggatgaa gatttagaaa tagtaaaaga
 2761 aagattaagc acatcaaaac gaataaataa tgaatgccca gaaacatcat caagaagtac
 2821 atcaagaagt acatcaaaaa gaataaatta taccactcca caaaaattaa tagaacaaaa
 2881 aatagaagaa ataaaccacac atcattatta tataacagga ataatggact aaagaaaata
 2941 ttaatacta ataaatacag ggcaggaaga gaattatgtt ataagagaac taataccaga
 3001 acaagagata gtaacaatag aacaacaaaa ttcagaacta ccaaagcgt taagaaaaaa
 3061 cgaagaaaca actgaaaaag aattaattat tggaggaata ccaatattaa taaattttaa
 3121 aatatatcaa ggagataaaa atattacact agggataaaa tggtagaaa aagtc aaacc
 3181 atataaatta gaagatagac aattaacaat aagttatgaa aataagaaaa taataataaa
 3241 aagaactttg atataatgcc aaagatatac atactttcca aaataatagt agaaggatat
 3301 tataatgat attatacacc tatggtggat acaggagcag aagctaatat gtgtagacat
 3361 aattgtttac cagaaagtaa atgggaaaag ctaaaaaacc ccatagtagt aacaggattt
 3421 aataatgaag gtagtatgat aacatataaa gcaagaaata taaaaatata aatatgggat
 3481 aaaatattaa ccatagaaga aatatatagt tacgaattca caaataaaga tatattatta
 3541 ggaatgccat ttttagataa attataccca catattataa caaaaacaca ttggtggtt
 3601 actaccctgt gtaacaaaaa attaggagca aaaagagtaa ataataaagt aagaaaaaca
 3661 acaccctgga ttaaaggaag tgaagagatt acccaaaaaa tagaaaaat aatacaagt
 3721 aaccataata tagagataat catttttca ataaataaga taaaccact acaagataaa
 3781 ctagaattac tatataatga taatccact caaggatggg aaaaacatca acaaaaaata
 3841 aagattgaac taatagatga aaatagcata ataacacaaa aacctttaaa atacaattt
 3901 aatgatttaa cagaatttaa aatgcatata aaagaattat tagataataa ctacatacaa
 3961 gaaagtaata gtaaacatac tagcccagca ttatagtaa ataagcatag tgaacaaaaa
 4021 agaggaaaaa gccgtatggt tatagattat agaaacttaa atgcaaaaac aaaaacatat
 4081 aattatccga taccaataaa aatactaaaa attagacaaa tacaaggata taactattt
 4141 agtaaatttg actataaatc aggattttac catttaaac tagaagatga atctaaaaag
 4201 ttaacagcat tcacagtacc acaaggattt tacgaatgga acgtattacc ttttgatat
 4261 aaaaatgcac caggtaggta tcaacattt atggataatt acttcaacca attagaaaa
 4321 tgtataatat atatagatga tatattgcta tattctagaa cagagaacga acatataaaa
 4381 ctactagaaa aatcatatac cattgtagaa atatcaggaa taagttaag taaaagaaa
 4441 gcagaagtaa tgaaaaatca aatagaattt ttaggatatac aaatggataa aaacggaata
 4501 aaaaatgcaaa cccatgtagt acaaaaaata attaaactga atgaacact tgatacaaaa
 4561 aagaagttac aatcattttt aggattggtt aaccaagtaa gagaatatat tcctaatta
 4621 gcagaaaact taaaccatt acagaaaaaa ttaaaaaagg acatagaata tcattttgac
 4681 gaaaaagata aaatacatat acagaagata aaaaatatgt gtaaaaaatt accaaaacta
 4741 tattttccag atgaaaagaa acaatttaca tatattgtag aaactgattc tagtgatcga
 4801 gtcgac

L8 ANSWER 135 OF 146 GENBANK.RTM. COPYRIGHT 2001
 LOCUS (LOC): NTA414164 GenBank (R)
 GenBank ACC. NO. (GBN): AJ414164
 CAS REGISTRY NO. (RN): 361143-47-7
 SEQUENCE LENGTH (SQL): 6060
 MOLECULE TYPE (CI): DNA; linear
 DIVISION CODE (CI): Plants, fungi, algae
 DATE (DATE): 3 Oct 2001
 DEFINITION (DEF): Nicotiana tabacum partial ORF2, ORF3 and ORF4,
 pararetrovirus-like sequence, V3.
 KEYWORDS (ST): movement protein; ORF2; ORF3; ORF4; polyprotein;
 translatio; transactivator/inclusion body protein
 SOURCE: common tobacco.
 ORGANISM (ORGN): Nicotiana tabacum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
 Tracheophyta; Spermatophyta; Magnoliophyta;
 eudicotyledons; core eudicots; Asteridae; euasterids I;
 Solanales; Solanaceae; Nicotiana

NUCLEIC ACID COUNT (NA): 2724 a 857 c 915 g 1559 t 5 others
REFERENCE: 1 (bases 1 to 6060)
AUTHOR (AU): Jakowitsch,J.; Mette,M.F.; van Der Winden,J.;
Matzke,M.A.; Matzke,A.J.
TITLE (TI): Integrated pararetroviral sequences define a unique
class of dispersed repetitive DNA in plants
JOURNAL (SO): Proc. Natl. Acad. Sci. U.S.A., 96 (23), 13241-13246
(1999)
OTHER SOURCE (OS): CA 132:103712
REFERENCE: 2 (bases 1 to 6060)
AUTHOR (AU): van der Winden,J.
TITLE (TI): Direct Submission
JOURNAL (SO): Submitted (24-SEP-2001) van der Winden J., Plant
Molecular Genetics, Institute of Molecular Biology,
Billrothstrasse 11, 5020 Salzburg, AUSTRIA

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..6060	/organism="Nicotiana tabacum"
		/db-xref="taxon:4097"
exon	<3..1157	/note="ORF2"
		/number=1
mRNA	<3..1157	/note="ORF2"
CDS	<3..1157	/note="ORF2"
		/codon-start=1
		/product="putative movement protein"
		/protein-id="CAC88780.1"
		/db-xref="GI:15963353"
		/translation="RLDLSIDKILEVPTIKNWFK RQKEEYYVVSQREHIIDCKYIKGK AQIPILNKRLLNKEIQDIKAKNPIKYVHLGGTEI LIKACFREGIDTPIEIYLADDRIV QPIEKSIISAVRGNLIYQKFKFIVSANYSVAVDD KNIDKSLVLYWKMSGIELAPGSKI FTARCKNLYVLTTHKITAKNKINKIKIENPFER IVSVIDNNDYSYTEIDMDEDLEIV KERLSTSKRINNEMPETSSRSTSRSKRYNTT PQKLIEQKIEEINPHHYITGIMD QRKYLILINTGQENYVIRELIPEQEIVTIEQQN SELPKALRKNEETTEKELIIGGIP ILINFKIYQGDKNITLGIKWLEKVKPYKLEDQRQL TISYENKKIHKRTLI"
mRNA	1156..2584	/note="ORF3 truncated ORF3"
exon	1156..2584	/note="ORF3"
		/number=1
CDS	1157..2584	/note="ORF3 aspartic proteinase, reverse transcriptase, ribonuclease H truncated"
		/codon-start=1
		/product="putative polypeptide"
		/protein-id="CAC88781.1"
		/db-xref="GI:15963354"
		/translation="MPKIYILSKIIVEGYNRY TPMVDGTGAETNMCRRHCLPESKWE"

KLKTPIVVTGFNNEGSMITYKARNIKIQIWDKIL
 TIEEISYEFTNKDILLGMPFLDK
 LYPHIITKTHWWFTTPCKQKLGAKRVNNKVRKTT
 PWIKGSEKITQKLENIQSNHNIE
 IIFSINKIKPLQDKLELLYNDNPLQGWKEHQTK
 IKIELIDENNIITQKPLKYNFNDL
 TEFKMHIELLDNNYIQESNSKHTSPAIVNKH
 EQKRGKSRMVIDYRNLNAKTKTYN
 YPIPKNILKIRQIQGYNYFSKFDCKSGFYHLKLE
 DESKKLTAFTVPQGFYEWNVLPFG
 YKNAPGRYQHFMDNYFNQLENCIIYIDDILLYSR
 TENEHIKLEKFIHIVEISGISLS
 KKKAUEVMKNQIEFLGIQIDKNGIKMQTHVVQKII
 NLNETLDTKKKLQSFLGLVNQVRE
 YIPKLAENLKPLQKKLKDIEYHFD"
 exon 3021..4229 /note="ORF4"
 /number=1
 mRNA 3021..4229 /note="ORF4"
 CDS 3021..4229 /note="ORF4"
 /codon-start=1
 /product="putative translation
 transactivator/inclusion body
 protein"
 /protein-id="CAC88782.1"
 /db-xref="GI:15963355"
 /translation="MLLQTTYQDKGTQDESQES
 KDIFTILTTLSQLMESMGKRLQQL
 ESQQHDYKNAELSRSEDSKLPKVEGDVGLQKTH
 NTVALYTAAGTSKQVSKKPHINVN
 LNTVFDKPFSTKKPREAIVAPQTSTYANSLHHN
 KKVYNHITQTYIENIYKIQTFLNL
 NPRSTTTTDDPTQDYVTQKLQGYNRLIAQPKTKAN
 LVKTCYNYGLLSTVYTHDGEEIIG
 IPELYKAFVTFKRITKGNLFFIKFYTTPAEILYD
 EIKPIQVIKIGLTRDMIPEEIE
 KQPEIQKMEIPSFYANKRIIGISTIIQELANNYL
 QGNAIWSYYSRDQLMIYANSKELR
 QGDMDEVQKWLSLLKPEMQPTTRALKKEFISNE
 LLTRYCKLVGHKYPDHICSKCNGD
 DNYVPEVQLE"
 repeat-unit 4282..4470 /rpt-type=TANDEM

SEQUENCE (SEQ):

1 gtcgactcga tctaagtata gacaaaatat tagaagtacc aacaataaag aattggttta
 61 aaagacaaaa agaagaatac tatgtagtaa gccaaagaga acatataata gattgtaaat
 121 acatcaaagg taaagcacia ataccaattt taaataaaag actactaaat aaagaaatac
 181 aagatataaa agcaaaaaat ccaataaaat atgtacactt aggaggaacg gagatcctaa
 241 taaaagcatg ttttagggaa ggaatagata ccctataga aatatacttg gcagatgata
 301 ggattgtaca acctatagaa aagagtataa taagtctgt aagaggtaac ttaatatacc
 361 taaaatttaa atttatagta agtgctaact attcagtagc agtagatgat aaaaatatag
 421 ataatcatt agtattatag tggaatgt ctggaataga attagcacca ggaagtataa
 481 tattcacagc aagatgtata aatctatag tcttaacaac gaaacataag ataacagcta
 541 aaaaataaat aaataaaata aaaatagaaa atccattcga aagaatagta tcagttag
 601 acaacaatga ttacagttag acagaaatag acatggatga agatttagaa atagtaaaag
 661 aaagattaag cacatcaaaa cgaataaata atgaaatgcc agaaacatca tcaagaagta
 721 catcaagaag tacatcaaaa agaataaatt ataccactcc acaaaaatta atagaacaaa

781 aaatagaaga aataaaccca catcattatt atataacagg aataatggac caaagaaaat
 841 atttaatact aataaataca gggcaggaag agaattatgt tataagagaa ctaataccag
 901 aacaagagat agtaacaata gaacaacaaa attcagaact accaaaagcg ttaagaaaaa
 961 acgaagaaaac aactgaaaaa gaattaatta ttggaggaat accaatatta ataaatttta
 1021 aaatataatca aggagataaa aatattacac tagggataaa atggttagaa aaagtcaaac
 1081 catataaatt agaagatagg caattaacaa taagttaga aaataagaaa ataataataa
 1141 aaagaacttt gatataatgc caaagatata catactttcc aaaataatag tagaaggata
 1201 ttataataga tattatacac ctatggttga tacaggagca gaaactaata tgtgtagaca
 1261 taattgttta ccagaaagta aatgggaaaa gctaaaaaca cccatagtag taacaggatt
 1321 taataatgaa ggaagtatga taacatataa agcaagaaat ataaaaatac aaatatggga
 1381 taaaatatta accatagaag aaatatatag ttacgaattc acaataaag atatatatt
 1441 aggaatgcca ttttagata aattataccc acatattata acaaaaacac atggtgtgtt
 1501 tactaccccg tgtaacaaa aattaggagc aaaaagagta aataataaag taagaaaaac
 1561 aacacccctgg attaaaggaa gtgaaaagat taccacaaaa ttagaaaaata taatacaag
 1621 taaccataat atagagataa tcattttctc aataataaag ataaaccac tacaagataa
 1681 actagaatta ctatataatg ataatccact ccaaggatgg gaaaacatc aaacaaaaat
 1741 aaagattgaa ctaatagatg aaaataacat aataacacaa aaacctttaa aatacaattt
 1801 taatgattta acagaattta aatgcataat aaaagaatta ttagataata actacataca
 1861 agaaagtaat agtaaacata ctageccagc atttatagta aataagcata gtgaacaaaa
 1921 aagaggaaaa agccgtatgg ttatagatta tagaaactta aatgcaaaaa caaaaacata
 1981 taattatccg ataccaataa aaatactaaa aattagacaa atacaaggat ataactattt
 2041 tagtaaatgt gactgtaaat caggatttta ccaattaaaa ctagaatg aatctaaaaa
 2101 gtaacagca ttcacagtac cacaaggatt ttacgaatgg aacgtattac ctttggata
 2161 taaaatgca ccaggtaggt atcaacattt tatggataat tactcaacc aattagaaaa
 2221 ttgtataata tatatagatg atatatgtct atattctaga acagagaacg aacatataaa
 2281 actactagaa aaattcatic acattgtaga aatacagga ataagtttaa gtaaaaagaa
 2341 agcagaagta atgaaaaatc aaatagaatt tttaggtata caaatagata aaaacggaat
 2401 aaaaatgcaa acccatgtag tacaaaaat aattaacttg aatgaaacac ttgatacaaa
 2461 aaagaagtta caatcatttt taggattggt taaccaagta agagaatata ttctaaatt
 2521 agcagaaaaac taaaaccat tacagaaaaa attaaaaag gacatagaat atcatttga
 2581 ctaaaaagat aaatacata tacagaagat aaaaaatag tgtaaaaaat tacaaaaact
 2641 atattttcca gatgaaaaga aacaatttac atatatgtga gaaactgatt ctagtatca
 2701 cagctacgga ggagtctaa aatataata tgataatgaa aaaattgaac accattgtag
 2761 atattattcc gcatcaca cggaaccaca attaaaatgg gaaataaata ggaagaact
 2821 atttgatta tataaatgtt ttttagcatt tgagccatat attgttata acaatttat
 2881 agtaagaaca gataatacac aagtaaaatg gtggataact cggaagtac aggattcagt
 2941 aactacaaag gaaataagga gacttgtatt aaatatacaa aattttacat ttacaattga
 3001 agtaatacga actgacaaga atgtattgac agactaccta tcaagacaaa ggtacccaaa
 3061 cagatgaaag ccaagaatca aaagacatat ttacaatcct tactactcta tccttacaga
 3121 tggagagtat gggaaaaaga ttacaacagc tagaaagtca gcagcatgac tataaaaatg
 3181 cggagctaag tgcagcgaa gactctaaac ttccaaaggt agaaggagac gttgggaaac
 3241 tccaaaaaac ccataacaca gttgctttat atagactgc aggtacaagc aaacaagtta
 3301 gcaaaaagcc acatatcaat gtaaacctaa ataccgtatt tgataagcca ttacatcaa
 3361 aaaagccaag agaagcaata gttatagccc cacaacttc aacctatgct aatagcctac
 3421 accacaacaa aaaggtatat aacctatata ctcaaacata tattgagaat atatatataa
 3481 tccagacatt tctgaacct aacctagat caacaactac tacagatcca acacaagatt
 3541 atgtaacca aaaaactacag ggatataata ggcttatagc ccagccaaaa acaaaagcaa
 3601 acttagtaaa aacatgttac aactacggac tactcagcac agtatatacc catgacggag
 3661 aagaaataat tggaaatacca gagctataca aagcatttgt caccttaag agaattacaa
 3721 aaggcaacct atctctac aaattctata caacaccagc ggaaatacta tatgatgaaa
 3781 taaaacccat aatacaggtg ataaaaatag gactaacagc agatatgata ataccggaag
 3841 agatagagaa acaaccggag atacagaaaa tggagatacc aagtttctat gccataaaaa
 3901 gaataattgg tatatcaact attattcaag aactagctaa caattatcta caaggaaatg
 3961 ctatctggag ctactatcc agagaccaat tgatgatata tgccaactcg aaggaaactac
 4021 gacaaggaga tatggatgaa gtccagaaat ggattttatc attattaaag ccagaaatgc
 4081 aaccaactac cagagcattg aaaaaagaat ttatttcaa cgagttatta acaagatact

4141 gcaaactagt aggacacaaa tatccagacc acatatgttc aaagtgaac ggagatgata
 4201 actatgtacc agaagtccaa ctagaatgaa ggtatcaaca agaagacaaa agaagaagac
 4261 agctactgga aacatataat gtaaagtaaa tagtactagt cgcattcatg aacagtaag
 4321 gtcgttcagt aatagtaaga gtcataaagt aaatagtata tgcataatc atgaacagta
 4381 aaggctgttc atgaatagta agagtcataa agtaaatagt atatgtcata atcatgaaca
 4441 gtaaaggctg tcatgaata gtaggagtc tttgtaaca gtaagagtcg ttttaattt
 4501 ctttatatag aatgtaaatc cgaggaaagga ggacatcctc agacatcctc atctcactc
 4561 tctctcttat ctctctcaa taaaatatct aattatatac aaacttcaga aagctatgga
 4621 gcattcaaac gagttacaaa accaggtgaag ttattcata atcatgttta actaaactca
 4681 tatattcctt ataactctct gaatatcata attgtttatc atgttatagt actgttataa
 4741 agaacatctt gagaagttt gcctgtctaa taatgtctgag agtagttaag cccagactat
 4801 gccatcctaa agtggaacc agtaggcaag gaagccgtt aggggagtag acagagttga
 4861 ggcgtgtta gggtaaatga ttgaagcatg aaaaacatgg tagtgaccag aaaagattgt
 4921 tcagtataac ttataaaat atgataaact ctatgataaa caaagaggtt agagggaata
 4981 tgtttagtaa acacatgata aggataaata gaaatctgaa tgaacaacca cactatttta
 5041 ttagaagaaa atattgacta caaaatactt atgttatata tctttaagag acttaataac
 5101 gatgttaaaa gaaaaattta cgatatttta gttacttttg aaataatata tgtaatggaa
 5161 caagcattta cggaactaat tgtatcacat gaaatagata tattgatctc atatgaacta
 5221 gatttatatt cagactaatg atcacatata aagttaataa aatcttttta taaatgaac
 5281 atagcacttt tcattatata agattctata ataccagaa atggaaactt cttaaaaccc
 5341 tcagtaacat aaatagaaaa atagattgtg ttaatacaa tattaacac tgcaagata
 5401 ttattagata taagctatga aaaataaaca ttatattaa gaagcaaatc atgaaaagc
 5461 aaaaatata gaagaatata aaaaactaa agatattaaa tacttagaat acacactga
 5521 aaatataaaa gaaattcaa taaaaaagat cctatgaaaa ataagatgaa ctcaacaac
 5581 cacccecaaa ttgacatac aaggatgact ttctagtgt gtacttat ccatctacn
 5641 ttggccagc ggagccatn tccactcaa tagaccctg gatcctcaac ttctacaaac
 5701 gatattgtat atggtaaaat ctaaaggctc aatttttgt cactatggca cttcataagc
 5761 acacttctat aggtctctaga ttatggcga gattaacacc tgagggggtta tcgacgccc
 5821 gcctcaagga agtcaggtt gattaagta tgattaaaac gtgggaggct ccaaggcatg
 5881 caactagaga tgacaaagtc tactaggcta gttcagacc gaacagggc attaaatgtt
 5941 ttaccatcc ntatatctt tcaataatgc attgtacta tgttgggatt cctctggtg
 6001 tataaaaggg gatcntgtc atttttaa aggatctgat gctcaatatt gaatacatag

L8 ANSWER 136 OF 146 GENBANK.RTM. COPYRIGHT 2001

LOCUS (LOC): NTA413172 GenBank (R)

GenBank ACC. NO. (GBN): AJ413172

CAS REGISTRY NO. (RN): 359759-08-3

SEQUENCE LENGTH (SQL): 6581

MOLECULE TYPE (CI): DNA; linear

DIVISION CODE (CI): Plants, fungi, algae

DATE (DATE): 3 Nov 2001

DEFINITION (DEF): Nicotiana tabacum pararetrovirus-like
sequence, V6.

KEYWORDS (ST): ORF2; ORF3; ORF4

SOURCE: common tobacco.

ORGANISM (ORGN): Nicotiana tabacum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
 Tracheophyta; Spermatophyta; Magnoliophyta;
 eudicotyledons; core eudicots; Asteridae; euasterids I;
 Solanales; Solanaceae; Nicotiana

NUCLEIC ACID COUNT (NA): 2893 a 897 c 1032 g 1759 t

REFERENCE: 1 (sites)

AUTHOR (AU): van der Winden, J.

JOURNAL (SO): Unpublished

REFERENCE: 2 (bases 1 to 6581)

AUTHOR (AU): van der Winden, J.

TITLE (TI): Direct Submission
JOURNAL (SO): Submitted (25-SEP-2001) van der Winden J., Plant
Molecular Genetics, Institute of Molecular Biology,
Billrothstrasse 11, 5020 Salzburg, AUSTRIA

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..6581	/organism="Nicotiana tabacum" /db-xref="taxon:4097"
mRNA	<1..1157	/note="ORF2"
CDS	<1..1157	/note="ORF2; putative movement protein" /codon-start=3 /product="hypothetical protein" /protein-id="CAC88194.1" /db-xref="GI:15799222" /translation="RLDLSIDKILEVPTIKNWFK RQKEEYYVVSQREHIIDCKYIKGK AQIPILNKRLNKEIQDIKAKNPIKYVHLGGTEI LIKACFREGIDTPIEYLAADDRIV QPIEKSIVSAVRGNLIYQKFKFIVSANYSVAVDD KNIDKSLVLYWKMSGIELAPGSKI FTARCKNLYVLTTHKHITAKNKINKIKIENPFER IVSVIDNNDYSYTEIDMDEDLEIV KERLSTSKRINNEMPETSSRSTSRSTSKRINYTT PQKLIEQKIEEINPHHYITGIMD QRKYLILINTGQENYVIRELIPEQEIVTIEQQN SELPKALRKNEETTEKELIIGGIP ILINFKIYQGDKNITLGIKWLKKVKPYKLEDRQL TISYENKKIIKRTLI"
mRNA	1157..1726	/note="ORF3"
CDS	1157..1726	/note="ORF3" /codon-start=1 /product="putative polyprotein (truncated)" /protein-id="CAC88195.1" /db-xref="GI:15799223" /translation="MPKIYILSKIIVEGYYNRY TPMVDTGAEANMCRHNCPLPESKWE KLKIPIVIGFNEGSMITYKARNIKIQIWDKIL TIEEIYSYEFTNKDILLGMPFLDK LYPHIITKTHWWFTTPCKQKLGAKRVNNKVRKTT PWIKGSEKITQKLENVIQSNHNE IIIFSINKIKPLQDKLELLYNDNLLQGWE"
mRNA	3021..3323	/note="ORF4"
CDS	3021..3323	/note="ORF4; putative translocation transactivator (truncated)" /codon-start=1 /product="hypothetical protein" /protein-id="CAC88196.1" /db-xref="GI:15799224" /translation="MLLQTTYQDKGTQTDSE KDIFTILTTLTSLQMESMGKRLQQL ESQQHDYKNAELSRSEDSKLPEVEGDVGKLQKPI"

TQLLYIQLQVQANKLAKSHISM"

repeat-unit 4281..4469 /rpt-type=TANDEM

SEQUENCE (SEQ):

```

1 gtcgactcga tctaagtata gacaaaaat tagaagtacc aacaataaag aattggttta
61 aaagacaaaa agaagaatac tatgtagtaa gccaaagaga acatataata gattgtaaat
121 acatcaaagg taaagcacia ataccaattt taaataaaag actactaaat aaagaatac
181 aagatataaa agcaaaaaat ccaataaaat atgtacactt aggaggaacg gagatcctaa
241 taaaagcatg ttttagggaa ggaatagata cccctataga aatatacttg gcagatgata
301 ggattgtaca acctatagaa aagagtatag taagtgtgtt aagaggtaac ttaatatacc
361 aaaaatttaa atttatagta agtgctaact atcagtagc agtagatgat aaaaatatag
421 ataatcatt agtattatac tggaaaatgt ctggaataga attagcacca ggaagtaaaa
481 tattcacgc aagatgtaa aatctatatg tcttaacaac aaaacataag ataacagcta
541 aaaaataaat aaataaaaa aaaaatagaa atccattcga aagaatagta tcagtatatg
601 acaacaatga ttacagtatt acagaaatag acatggatga agatttgaa atagtaaaag
661 aaagattaag cacatcaaaa cgaataaata atgaaatgcc agaaacatca tcaagaagta
721 catcaagaag tacatcaaaa agaataaatt ataccactcc acaaaaatta atagaacaaa
781 aaatagaaga aataaaccca catcattatt atataacagg aataatggac caaagaaaat
841 atttaact aataaatata gggcaggaag agaattatgt tataagagaa ctaataccag
901 aacaagagat agtaacaata gaacaacaaa atcagaact accaaaagcg ttaagaaaaa
961 acgaagaac aactgaaaa gaattaatta ttggaggat accaatatta ataaatttta
1021 aaatatatca aggagataaa aatattacac tagggataaa atggttaaaa aaagtcaaac
1081 catataaatt agaagatagg caattaacaa taagttaga aaataagaaa ataataataa
1141 aaagaacttt gatataatgc caaagatata catactttcc aaaataatag tagaaggata
1201 ttataataga tattatacac ctatggtgga tacaggagca gaagctaata tgttagaca
1261 taattgttta ccagaaagta aatgggaaaa gctaaaaatc cccatagtag taataggatt
1321 taataatgaa ggaagtatga taacataaa agcaagaat ataaaaatc aaatatggga
1381 taaaatatta accatagaag aaatatatag ttacgaattc acaataaag atatatatt
1441 aggaatgcca ttttagata aattatacc acatattata acaaaaac attggtggtt
1501 tactaccccg tgtaacaaa aattaggagc aaaaagagta aataataaag taagaaaaac
1561 aacaccctgg attaaaggaa gtgaaaagat taccacaaaa tagaaaatg taatacaaaag
1621 taaccataat atagagataa tcatcttc aataataaag ataaaaccac tacaagataa
1681 actagaatta ctatataatg ataactact ccaaggatgg gaataacatc aaacaaaaat
1741 aaagattgaa ctaatatagtg aaaatagcat aataacacaa aaaccttaa aatacaattt
1801 taatgattta acagaattta aatgcataa aaaagaatta ttagataata actacataca
1861 agaaagtaat agtaaacata ctaagccagc atttatagta aataagcata gtgaacaaaa
1921 aagaggaaaa agccgtatgg ttatagatta tagaaacta aatgcaaaaa caaaaacata
1981 taattatccg ataccataa aataactgaa aattagacaa atacaaggat ataactattt
2041 tagtaaatgt gactgtaaat cgggatttta ccatthaaaa ctagaagatg aatctaaaaa
2101 gttacagca ttcacagtac cacaaggatt ttacgaatgg aacgtattac cttttggata
2161 taaaatgca ccaggtaggt atcaacattt tatggataat tacttcaacc aattagaaaa
2221 ttgtataata tatatagatg atatatgtct atattctaga acagagaacg aacatataaa
2281 actactagaa aaattcatc acattgtaga aatatcagga ataagtttaa gtaaaaagaa
2341 agcagaagta atgaaaaatc aaatagaatt tttaggata caaatagata aaacggat
2401 aaaaatgcaa acccatgtag tacaaaaat aattaacttg aatgaacac ttgatacaaa
2461 aaagaagta caatcatttt taggattggt taaccaagta agagaatata ttctaaatt
2521 agcagaaaac taaaaccat tacagaaaaa attaaaaaag gacatagaat atcattttga
2581 cgaaaaagat aaaatcataa tacagaagat aaaaaaatg tgtaaaaaat taccaaaact
2641 atattttcca gatgaaaaga aacaatttac atatatgtga gaaactgatt ctagtgtaca
2701 cagctacgga ggagtcttaa aatataata tgataatgaa aaaaatgaac accattgtag
2761 atattttcc ggatcataca cggaaccaca attaaaatgg gaaataata ggaaagaact
2821 atttgatta tataaatgtt ttttagcatt tgagccatat attgtttata acaaatattt
2881 agtaagaaca gataatacac aagtaaaatg gtggataact cggaagtac aggttcagt
2941 aactacaag gaaataagga gactgtattt aaatatacaa aattttacat ttacaattga
3001 agtaatacga actgacaaga atgttattgc agactaccta tcaagacaaa ggtacccaaa
3061 cagatgaag ccaagaatca aaagacatat ttacaatcct tactactcta tcttacaga

```

3121 tggagagtat gggaaaaaga ttacaacagc tagaaagtca gcagcatgac tataaaaatg
 3181 cggagctaa gctatcgga gactctaaac ttccagaggt agaaggagac gttgggaaac
 3241 tccaaaaaac cataacacag ttgctttata tacagctgca ggtacaagca aacaagttag
 3301 caaaaagcca catatcaatg taaacctaaa taccgtattt gataagccat ttacatcaaa
 3361 aaagccaaga gaagcaatag ttatagcccc acaaaactca acctatgcta atagcctaca
 3421 ccacaacaaa aaggtatata accatattac tcaaacatat attgagaata tatataaaat
 3481 ccagacattt ctgaacctaa accctagatc acaactact acagatccaa cacaagatta
 3541 tgaacccaa aaactacagg gatataatag gcttatagcc cagccaaaaa caaaagcaaa
 3601 cttagtaaaa acatgttaca actacggact acttagcaca gtatataccc atgacggaga
 3661 agaaataatt ggaataccag agctatacaa agcatttgc acctcaaga gaattacaaa
 3721 aggcaacta ttcttcata aattctatc agcaccagcg gaaatactat atgatgaaat
 3781 aaaaccata atacaggtga taaagatagg actaacacga gatatgataa taccagaaga
 3841 gatagagaaa caaccggaga tacagaaaat ggagatacca agttctatg ccaataaaag
 3901 aataattggt atagcaacta ttattcaaga actagctaac aattatctac aaggaaatgc
 3961 tatctggagc tactattcca gagaccaatt gatgatat gccaaactga aggaactacg
 4021 acaaggagat atggatgaag tccagaaatg gattttatca ttattaaagc cagaaatgca
 4081 accaactacc agagcattga aaaaagaatt tattcaaac gagttattaa caagatactg
 4141 caaactagta ggacacaaat atccagacca catatgttca aagtgcacg gagatgataa
 4201 ctatgtacca gaagtccaac tagaatgaag gtatcaaaa gaagacaaa gaagaagaca
 4261 gctactggaa acatataatg taaagtaaat agtactagtc gcattcatga acagtaaagg
 4321 tcgttcata atagtaagag tcataaagta aatagtatat gtcataatca tgaacagtaa
 4381 aggtcgttca tgaatagtaa gactcataa gtaaatagta tatgtcataa tcatgaacag
 4441 taaaggtcgt tcatgaatag taggagtcgt ttgtaaacag gaagagtcgt ttaattttc
 4501 ttatataga atgtaaatct gaggaaggag gacatctca gacatctca cccctacctt
 4561 ctctctctaa tattctctct tatctctct cttgtaagaa atatctgagt taattacaag
 4621 tatctaaaac agctatggag cattcaaacg agttacaaa ccaggtaagt tattcataa
 4681 tcatgtttaa ctaaactcat atattcttta taatctctg aatatcataa ttgtttatca
 4741 tgttatagta ctgtttataa gaacatctg agaagtttg cctgtctaat aatgctgaga
 4801 gtatgttaag ccagactatg ccatactaaa gtggaacca gtaggcaagg aagccgttta
 4861 ggggagtata cagagttgag gcgctgttag ggtaaatgat tgaagcatga aaaacatggt
 4921 agtgaccaga aaagattggt cagtataact tattaaaata tgataaactc tatgataaac
 4981 aaagaggtta gaggaatat gtttagtaaa cacatgataa ggataaatag aaatctgaat
 5041 gaacaaccac acgtatttat tagaagaaa tattgactac aaaatactta tgttatatat
 5101 cttaagaga cttaataacg atgttaaaag aaaaatttac gatattttag ttacttttga
 5161 aataatatat gtaatggaac aagcatttac ggaactaatt gtatcacatg aaatagatat
 5221 attagatcta tatgaactag atttatattc agactaatga tcacatatca agttaataaa
 5281 atctttttat aatgaaaca taacactttt cattatataa gattctataa taccagaaa
 5341 tggaaacttc ttaaaaacct cagtaacata aatagaaaaa tagattgtgt taataacaat
 5401 attaaaacct gcaaagatat tattagatat aataaattac gtaataaagc ttactaact
 5461 atagaaatg aaattgaatt ttatgaagat aaacttgaag gttatcaaca cagaaaagaa
 5521 ataacactga gaaatataa aacctggat atatgtatca acaggtatc atctcagat
 5581 taatatgaac aatcaaaaa ttgatatgta caaatattt aagtactggt tagaaattct
 5641 aaaacatata aatatgtcaa aaataatatt agaaaaaat atagaagaat atcaaaaaac
 5701 taaagatatt aataacttag aatacacact tgaaaatata aaagaaattt caaggttaat
 5761 ttcgctagct acagataatc atgaaaaagc atttaattcc acaactacat aaaaaagatc
 5821 ctatgaaaaa taagatgaac tcaacaacc ccattctggt ggggatctaa tgatatatgg
 5881 attggtgaag gtaagaagga gcggatggct ttgtatctaa gatcacaagg tataggcttg
 5941 gacaattatt ggtttgtata ttggtatgta tcatacatc acaagtttg gtatagaagt
 6001 tgcagtcagg ataaatatag cactcatagc tattctgtaa ttctaattg catgtgtaa
 6061 ctgggatcta atagcagtag aggtgctagt atctgtttt attttctag gtattgtcat
 6121 ggcaccatta agtggattg gtatgtttac atgcccaccc tagagctgtc aaggtagtgc
 6181 atggctatgg ttgaaatgc aggtattggc atattgctc tggaggcagc tgtgtcatt
 6241 ttctcaact ctagacgagt ttaaattgat gtcaagagtt tctgggcaa ttgtcagtc
 6301 attcgttgat ccagggaaga agaaatagtc atgcccaccc atagattttg cacaattgtg
 6361 ctactttaat attgggacac ttccaaaagt tgatggtatt actaggagat tcacatagc
 6421 cttgtctttt ataagcaaa acatggcgca tggtagaga tttagagctg ctacattgag

6481 gtcttgccca tggcttggg atacacatac caccatatgt ttggagtgt ctatgatac
6541 atttgaaca gttagtcta ttacatagtt agttcaagct t

L8 ANSWER 137 OF 146 GENBANK.RTM. COPYRIGHT 2001

LOCUS (LOC): AF190123 GenBank (R)

GenBank ACC. NO. (GBN): AF190123

CAS REGISTRY NO. (RN): 248236-12-6

SEQUENCE LENGTH (SQL): 7767

MOLECULE TYPE (CI): DNA; linear

DIVISION CODE (CI): Viruses

DATE (DATE): 23 Jun 2000

DEFINITION (DEF): Tobacco vein-clearing virus, complete genome.

SOURCE: tobacco vein-clearing virus.

ORGANISM (ORGN): tobacco vein-clearing virus

Viruses; Retroid viruses

NUCLEIC ACID COUNT (NA): 3729 a 959 c 1150 g 1929 t

REFERENCE: 1 (bases 1 to 7767)

AUTHOR (AU): Lockhart,B.E.; Menke,J.; Dahal,G.; Olszewski,N.E.

TITLE (TI): Characterization and genomic analysis of tobacco vein
clearing virus, a plant pararetrovirus that
is transmitted vertically and related to sequences
integrated in the host genome

JOURNAL (SO): J. Gen. Virol., 81 Pt 6, 1579-1585 (2000)

OTHER SOURCE (OS): CA 133:306131

REFERENCE: 2 (bases 1 to 7767)

AUTHOR (AU): Lockhart,B.E.; Menke,J.; Dahal,G.; Olszewski,N.

TITLE (TI): Direct Submission

JOURNAL (SO): Submitted (25-SEP-1999) Plant Biology, University of
Minnesota, 220 Biological Sciences Center, St. Paul, MN
55108, USA

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..7767	/organism="tobacco vein-clearing virus" /specific-host="Nicotiana edwardsonii" /db-xref="taxon:107324"
CDS	416..2155	/note="ORF 1" /codon-start=1 /product="putative coat protein" /protein-id="AAF08287.1" /db-xref="GI:6425073" /translation="MNKEEFAIDEKTYENQDGLK IKIIFSNLGRRYKKIGDQIYLMIE KETARLEDSLAMTRIIKENEEIDKKREIEIIRN QANKEIQQLEETKNTKIIALEKEL NMLKLLYENKQREKDKEIELTDEINKFKQKLQPE ENLRIIEIDNNIDDQKSELSEISE TYTEILENIEKTKNLEINTGDVNMDEKPSTSGIK NPKELNPSYYTVSYEQSDRNNTLW NSRLNKKWTPKPIHEQYNFLDLDCVEDINKTIQL WIGYISKQLIDNKIGITETPGYIE RTLIGTVKLWLHNLTKESIDTLRSNKKFNGDMAT TSMEILYKYEIAIRNEFSSMTTET


```

EEQHKEKQINRNLMTKLAICNMCYIDEYTCAFRE
YYYKGTYNTEEGREIRKLYFTKLP
EPFSSKIIKDWNEAGLDTLGARIKFLQQWVQL
CEKYKEEIKMEKILIKNLKCKNK
TAPQFGCTDKYNKKWKTKKYKKHRSKYKYKKPRK
RYYVKNYQAKKPFPRPKKLTECTC
YNCGKLGHIKDKAPKNPKKKQITEIIDDEEY
MQMEYIDYELDSSEDSIYEISDSDD
EIDSEITEEEDI"
CDS      2155..3414      /note="ORF 2"
                        /codon-start=1
                        /product="putative cell-to-cell
                        movement protein"
                        /protein-id="AAF08288.1"
                        /db-xref="GI:6425074"
                        /translation="MTEKEIKVVAREEYQNEESS
EQKIIFDNNIFEQIKGKELDLSVE
KIFEVPTIKNWFKRQKEEYVVSQKEHIIDCKYT
GGKAKIPIINKRIINKEIQDIKAK
NPIKYVHLGGTEILIKACFREGIDTPIEYILADD
RIIQPIEKSIISAIKGNLIYQKFK
FIISANYSVALNDRNIDKSLVLYWKMSGIELAPG
SKVFTARCKNLYVLTTKHKIAAKN
KINKIKIENPFERIVTVIDNNEYSYKEIDMEEDL
EIVKERLSTSSIPKQLTHETPTSS
RISTSKREYIIPKNLARNTIEEIKPYHYITGIM
EQRKYQILINTGQENYIIRELVA
ESEIITTEQLCPKELIINEEITEKEIIGGI
PLIIQFKIYQGNENITLGKWLER
LKPYNLEDKQLTITYKNKKVIIKRTTE"
CDS      3411..5318      /note="ORF 3; contains aspartic
                        protease, reverse transcriptase
                        and RNase H"
                        /codon-start=1
                        /product="polyprotein"
                        /protein-id="AAF08289.1"
                        /db-xref="GI:6425075"
                        /translation="MKIYILAKIIVEGYYNRYYT
PMIDTGAEANICKYNCLPTDKWEK
LKTPMVVTGFNNEGSMINYKARNVKIQIWDKILT
IEEIYNFEFTTKDMLLGMPFLEKL
YPHIITKTHWWFTTPCKNKVGAKRVNNKQRKTTE
WIRGSEKITQKLENINKNTTTQLE
IIIFTIDKVKIIQKKLEKLYNDNPLQGWEKHKTK
VKIELIEENSIITQKPLKYNFNDL
TEFKIHIKDLLDNKYIQESNSKHTSPAFIVNKHS
EQKRGKSRMVIDYRNLNAKTKTYN
YIPNKILKIRQIQGYNYFSKFDCKSGFYHLKLE
DESKKLTAFVTPQGFYEWNVLPFG
YKNAPGRYQHFMDNYFNQLENCIVYIDDILLYSR
TQDEHIKLLKFAHIIENSGISLS
KTKAEIMKNQIEFLGIQIDKNGIKMQTHIVQKII
NLDENIDTKKKLQSFLGIVNQVRE
YIPKLAENLKPLQKKLKKDVEYSFDEKDKEQIKK
IKILCKKLPKLYFPDENKKFTYIV
ETDSSNYSYGGVLKYRYNKEKIEHHCRYYSYSY

```

EPQEKWEINRKELFALYKCLLAFE
 PYIVYTRFIVRTDNTQVKWWITRKVQDSVTTKEI
 RRLVLNLTFTTIEINTKNV
 ADYLSRQSYPN"
 CDS 5275..6528 /note="ORF 4"
 /codon-start=1
 /product="putative transactivator
 factor"
 /protein-id="AAF08290.1"
 /db-xref="GI:6425076"
 /translation="MLLQIYQDKATQTDPDNTM
 ENLLAMTTLCKKVESMDEEIQIM
 RKTASSQQHDSKNAELRQSEVSKIPELEGDVEKH
 LKTHNGLLNTVAGSSTTSSSSARK
 KEDIKPRYTNINMNNLFSKPFVQRNTQDTQKEIF
 LPPQVNTYKESLNQTKKTYNHITR
 TYIDNIHKIQNFLNKNPRSQSTQNPEDYITHHL
 TGYNKLIALPNTSAKL VATCYNYG
 LLDTVYTQTGQEIANIPELYKAFMQYKRITKGT
 FYVRFYSATAEILYEEIKPIQVI
 KIGLTREMIPEKIEEQEEIEKVNIPEFYANKRI
 IGISTILNELANNYLNQNAIWSYY
 SREQTMIYSNCRDIREADMEELRQWVLSLLKPEQ
 PTTTRAIRRNFI SPDLLTRYCKLI
 SHKYPDHLCSKCKGEDNIVPDVQLE"

SEQUENCE (SEQ):

1 tggatcaga gccagtacac aaaaagaac acatcaagtt aataaagaag accaaataag
 61 aaacagtga agaaaaacaa acaagaaact aagaaggtta tactctgaat ataaaaaatt
 121 acacataaca agatagata ataataaggtt agaacaacta atagatatca tatctaatac
 181 agaataaaa tatatccagt atttgaaaa acagtatgaa caaggaagaa tataaacaac
 241 aactttttga aaaaattaca gaaaaaagac aagaattacg acatagacaa gataggctta
 301 agaataatat atttgctgga acttgccaga aatctttact agtcaaaga aaaattatat
 361 taaatttaca aatacaata caacagtttg aatatatact ccaacataac ctataatgaa
 421 taaagaagaa ttgcatag atgaaaagac atagaaaac caagatggat tgaataaaa
 481 gataatatt tcaatctag gtagaagata taaaaaata ggagaccaa tataccta
 541 gatagaaaa gaaacagcaa gattagaaga tagcctgaca gccatgacaa gaataattaa
 601 ggaatgaa gaaattgata aaaagaggga aattgaaatt atcagaaatc aagcaataa
 661 ggaatacaa caattggaag aaaccaagaa taaaaaatt atagcattag aaaaagaatt
 721 aaatatgta aattactat atgaaaaaa acaagagag aaagataaag aaatagaatt
 781 aacggatgag ataataaat tcaacaaaa attacaacca gaagaaaacc tcagaattga
 841 agaaatagat aataatatag atgacaaaa gtcagagcta agcgaataa gtgaacata
 901 tacagaaatt ttagaaaaa tagaaaaaac aaaaaattta gaaataata caggggatgt
 961 aaatatggat gaaaaaccta gcacatcagg tataaagaat caaaagaat taaaccaag
 1021 ttattataca gtaggttatg acaatctga tagaaataat acactatgga atagtagact
 1081 aaataaaaaa tggacaccaa aaccaataca tgagcaatat aacttttag atttagattg
 1141 ttagaggat ataataaaa caatccaact atggatagga tatatatcta acaataat
 1201 agataataaa ataggaataa cagaacacc aggatataa gaaaggacac ttataggac
 1261 agtaaaaatta tggttacata atttaacaaa agaaagtata gatactttaa gaagtaataa
 1321 aaaatttaatt ggtgacatgg caacaacgag tatggaata ttatacaaat atgaaatagc
 1381 cataagaac gaatttagta gtatgactac agaaacagaa gaacaacata agggaaaaa
 1441 aataataga aacctaata caaaattagc aatatgta atgtgctata tagatgaata
 1501 tactgtgca tttagagaat attattataa aggaacatat aatacagaag aaggcagaga
 1561 aataagaaaa ttatattca caaaattacc agaacccttt agtcaaaaa taataagga
 1621 ttggaatgaa gcaggataa cagatacttt aggagctagg ataaaaatt tacaacaatg
 1681 gttgtacaa ttatgtgaga aatataaga agaaataaaa atggaaaaaa tattaataa

1741 aaacttaaaa tgttgtaaaa ataaaacagc accacaattt ggatgtacag ataaatataa
1801 caaaaaatgg aaaacaaaga aatataaaaa acataggtca aagtataaat ataaaaaacc
1861 aagaaaaaga tattatgtaa aaaattatca ggctaaaaaa ccatttagac caaaaaagaa
1921 actaacagaa tgtactgtt ataactgtgg aaaattagga catatagcta aagactgcaa
1981 agcacctaaa aacccaaaga aaaaacaaat aactgaaatt ataattgatg atgaagaata
2041 tatgcagatg gaatatatag attatgaact agatagttaa gacagtatat atgaatatc
2101 agattcagat gatgagatag atagtgaat aactgaagaa gaagatgaaa tctaatgact
2161 gaaaaagaaa taaaagtagt agcacgagaa gaatatcaaa atgaagaatc ctcagaacaa
2221 aaaattatat ttgataataa tatattgaa caaataaaag gaaaaaatt agaccttagt
2281 gtcgaaaaaa tatttgaagt acctacgata aaaaattggt tcaaaagaca aaaagaagaa
2341 tactacgtag ttgccaataa agaactata attgattgca aatatacagg aggttaaggcc
2401 aaaataccaa taataacaa aagaataata aataaagaaa tacaagatat aaaggcaaaa
2461 aacccaatta agtatgtaca cttaggagga actgaaatat taataaaagc atgttttaga
2521 gaaggaatag ataccctat agaatatat ctacgagatg atagaataat ccaacctata
2581 gaaaaaagca taataagtgc aataaaagga aatcttatat atcaaaagt taaattata
2641 ataagcgcca actattcagt agcactaaat gatagaata tagataaatc attagtatta
2701 tactggaaga tgcaggaat agagttagca cccggaagta aagtatttac agccagatgt
2761 aagaacctat acgtattaac aacaaaacat aaatagcag cgaataataa aattaataa
2821 atcaaatag aaaacccgtt tgaagaata gttacagtta ttgacaataa tgaatatagc
2881 tataaagaaa ttgacatgga agaagatta gaaatagtaa aagaagatt aagtacatca
2941 agtatacca aacaattaac tcatgaaaca ccaacatcat caagaataag tacatcgaaa
3001 agagaatata taatccaaa aaatttagcc agaacacaa tagaagaat aaagccatat
3061 cattattata taactggaat tatggaaca cgaataatc aaatattaat aaacacagga
3121 caggaagaaa attatatcat aagagaatta gtacgagaaa gtgaataat aactactgaa
3181 caactatgcc ctgaattacc taaagaatta ataataatg aagaataatc agaaaaagaa
3241 ataattatg gaggaatacc tctaataa caatttaaaa tatatcaagg aaacgaaaat
3301 attacactag gaataaaatg gttagaaga ttaaaccat ataattaga agataagcaa
3361 ttaactataa catataagaa taaaaagta ataataaaaa gaacagaaga atgaaatat
3421 acatactgc aaaaatcata gtgaaggat attataatag atactataca ccaatgatag
3481 atacaggagc agaagcta atagttaaat ataattgtt accaacagat aaatgggaaa
3541 aataaaaaac acctatggt gtaacaggat ttaataatga aggtagtatg attactaca
3601 aagctagaaa cgtaaaaaa caaatatggg ataaaatatt aactatagaa gaaatatata
3661 accttaggt taccacaaa gatagtgt taggaatgcc attctagaa aagtataacc
3721 cacatataat acaaaaaa cactggtggt ttaccacacc atgtaaaaat aaagtaggag
3781 ctaaaagagt aaataataa caacgaaaa ctacagaatg gatacagggc agtgaaaaga
3841 tcacacaaa attagaaaat ataataaaa acacaactac ccaattagaa attatcatat
3901 ttacaataga taaagtaaag ataaccaga aaaaattaga aaaattatat aatgataacc
3961 ctctacaagg atgggagaaa cataaaaca aagtaaaaat agaactaatt gaagagaata
4021 gtataatac acaaaaacct ctaaaata actttaatga ttaacagaa ttaaaaatac
4081 atataaaga tttattggt aataatata tacaagaag taatagcaa catacaagtc
4141 cagcattat agtaataaa catagtgaac aaaaagagg aaaaagcagg atggtcatag
4201 actatagaaa ttaaatgca aaactaaaa catataata tccaatacc aataaataat
4261 taaagataag acaaatataa ggatataatt actttagtaa attgactgt aagtcaggat
4321 tttatcatt gaaactagag gatgaatca aaaaattaac agctttaca gtaccacaag
4381 gatttatga atggaatgt ttaccattg gatataaaa tgcaccagga agataccaac
4441 attttatga caactattt aatcaactag aaaattgtat agtatatata gatgatatac
4501 tactatattc aagaacaaa gatgaacata taaaattatt agaaaaatc gcacatatta
4561 tagaaaactc tggtaagc ttaagcaaaa ccaagcaga aattatgaaa aatcagatag
4621 aatttttagg aatcaaaata gataaaatg gaataaaaat gcaactcat atagtacaaa
4681 aaataatcaa tcttgatgaa aacattgata caaaaaagaa attacaatca tttctaggaa
4741 tagtaacca agtaaggga tatataccta aattagcaga aaacttaaaa ctttacaga
4801 aaaaactaaa aaaggatgta gaatatagt ttgatgaaa agataaggaa caaataaaga
4861 aaattaagat attgtgtaa aaattacca aattatact tccagatgaa acaaaaaaat
4921 ttacctatat ttagaaca gattcaagta attatagcta tggaggagt ctttaagata
4981 gatataata agaaaaata gaacatcatt gcagatact ttcaggatct tatacagaac
5041 cacaagaaa atgggaata aatagaaaag aattattgc attataaaa tgtctattag

5101 cattgaacc atatagatt tacaccagat ttattgtaag aacagacaat acacaagtaa
 5161 aatggtggat aactagaaaa gtacaagatt cagtcacaac aaaagaaata cggagactgg
 5221 tattaaatat attaaacttt acatttaca ttgaaataat taactaac aagaatgtg
 5281 ttgcagatta ttatcaaga caaagctacc caaactgac cagataatac aatggaaaat
 5341 ctactcttgg ccatgactac actttgtaaa aaagtggaaa gtatggatga agagatacaa
 5401 ataagagaa aaacagctag cagtcagcag catgactcaa aaaatgcgga gctaagacaa
 5461 tcggaagtct ctaaaattcc agagctagaa ggtgacgtg agaaacacct aaaaactcat
 5521 aacggtttgt taaatacagt tgcaggaagc agcacaacta gcagttcatc tgcaaggaaa
 5581 aaggaagata ttaaaccctag atactaat ataaatatga acaatttatt tcaaaaacca
 5641 ttctacaga gaaataccca agatacccag aaagaaatat tcttaccacc acaggtaaat
 5701 acctacaaag aaagtttaaa ccaaaccaaa aagacatata accatataac cgtacctat
 5761 atagacaata tacataaaat acagaatttt ttaataaaa atccaagatc ccaatctacc
 5821 caaaatccaa acgaagacta tattaccac catctaacag gatacaataa attgatagct
 5881 ttaccaata caagtgcata attagtagcc acctgtaca actacggact actagatact
 5941 gtatatcac agacaggaca agagatagcc aatatacctg agttatacaa ggcattcatg
 6001 caatataaga ggataactaa agggacatta ttctactac gattctattc ggctacagca
 6061 gagatactat atgaagagat aaagcctatc atacaagtta tcaaaattgg atcacaagg
 6121 gaaatgatca taccagaaaa gatagaagaa caagaagaga tagaaaaagt aaatattcca
 6181 gaattctatg ctaacaaag gataattgga atattacta tactaaatga actagcaaac
 6241 aactatctaa atcagaatgc aatttgagc tattattcaa gggaacagac aatgatatat
 6301 tcaaaattga gagatataag agaagcagat atggaagaac taagacagtg ggtcttaagt
 6361 cttctaaac cagaacaacc tacaaccaca agagctataa ggaggaattt catctccca
 6421 gatctgttaa caagatactg caagctaatt agtcacaaat atccagacca cctatgtca
 6481 aaatgcaaag gagaagataa cattgttcca gacgtacaac tggataaatt cagatagaag
 6541 aaaacgacaa aacaacaaga tatgacagaa agatctactt ttatagattc cttttctt
 6601 tgtaagaaa tgtattttg tttttataa aagtcgtaa gaatagtta ctttataaa
 6661 gtaagagta tcatgaaca gtaaggtcg tcatgaata gtaagagtcg ttgtataa
 6721 gtaagagtcg ttgtataa gtaagagtcg ttgtataa gtaagagtcg ttgtataa
 6781 tcgtttgtga atagtaagag tcaattttg tattataat agcagttcaa atgtgaataa
 6841 aaaacaggct gcagtttca agcatccaac aattcctctc tcttctctc aatattttt
 6901 gcagatataa gcatacga gctatggaac aacaaggaga tagcagaat ctacaaaaac
 6961 aggtatatca tttccaact atgaacagct aaattactat attcctgata atttctatgt
 7021 aaattataat tgttcatcat actatagtag tgtataaaa atcatcttga gaaagtttgc
 7081 catgataata tgctaatagt aggtaaacc agactatgcc atcctaattg tgaaccgggt
 7141 aggcagagaa gccgtttagg ggagtaaacc aagttgaagg cgctgttagg gtaattgatc
 7201 gaagcagaaa atcatgtag tagaccagaa agatgactaa aataacttat tataatatga
 7261 tgaataagta taatagacat agagattaga aggaatatgt ttagcaaatc atatgaaaa
 7321 tgaatactta ctattagat gatgatact gttataagtt actcatactata tatatctaa
 7381 aaaatatcga aaatgataca agaagaataa tctatgaaa actagtact ttgaaatgt
 7441 tagaaataat aaatcaaaat tggacagaat taattatacc agatgatata aatagact
 7501 tatatgagtt agacttatat tcagaataat aatgaatcat atataccaac aatattgca
 7561 acaaaactact gaaagcataa ctttactagg aaatttagta aaaacaaata tagaagagta
 7621 tcaagaacc caagatatac aatatataga atacatttg gaaagcttga aagaataact
 7681 tagattaatt ccgtacaaa aagaatatta tgaaaaagca ttaatttaa ataattaaat
 7741 gacaaaagat aagaccaccc ccaatt

L8 ANSWER 138 OF 146 GENBANK.RTM. COPYRIGHT 2001
 LOCUS (LOC): NTA238747 GenBank (R)
 GenBank ACC. NO. (GBN): AJ238747
 CAS REGISTRY NO. (RN): 225689-87-2
 SEQUENCE LENGTH (SQL): 7981
 MOLECULE TYPE (CI): DNA; linear
 DIVISION CODE (CI): Plants, fungi, algae
 DATE (DATE): 11 Nov 1999
 DEFINITION (DEF): Nicotiana tabacum pararetrovirus-like
 sequence, ORF1, ORF2, ORF3 and ORF4.

KEYWORDS (ST): coat protein; inclusion body protein; movement protein;
polyprotein; translation transactivator

SOURCE: common tobacco.

ORGANISM (ORGN): Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta;
Tracheophyta; Spermatophyta; Magnoliophyta;
eudicotyledons; core eudicots; Asteridae; euasterids I;
Solanales; Solanaceae; Nicotiana

NUCLEIC ACID COUNT (NA): 3814 a 1023 c 1189 g 1955 t

REFERENCE: 1 (bases 1 to 7981)

AUTHOR (AU): Jakowitsch, J.; Mette, M.F.; van Der Winden, J.;
Matzke, M.A.; Matzke, A.J.

TITLE (TI): Integrated pararetroviral sequences define a unique
class of dispersed repetitive DNA in plants

JOURNAL (SO): Proc. Natl. Acad. Sci. U.S.A., 96 (23), 13241-13246
(1999)

OTHER SOURCE (OS): CA 132:103712

REFERENCE: 2 (bases 1 to 7981)

AUTHOR (AU): Jakowitsch, J.

TITLE (TI): Direct Submission

JOURNAL (SO): Submitted (16-APR-1999) Jakowitsch J., Institute of
Molecular Biology, Austrian Academy of Sciences,
Billrothstrasse 11, Salzburg, Salzburg, A - 5020,
AUSTRIA

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..7981	/organism="Nicotiana tabacum" /cultivar="petit havana SR1" /db-xref="taxon:4097"
primer-bind	1..12	/partial /gene="tRNA-Met" /note="putative minus-strand DNA primer"
gene	1..12	/gene="tRNA-Met"
CDS	267..2063	/note="ORF1" /codon-start=1 /product="putative coat protein" /protein-id="CAB42620.1" /db-xref="GI:4775494" /translation="MNKEEFALEEKTYENPEGLK ITIIFSNLGRRYKKIGNNLNMLE KETVKLEDSLAMVRLTKENEEIDRKREIKEIQQ QAKEKIQQIEEVKNTKITELEKEL EMLKQMYANKLKEKEKRKEEEEELKLTNEIERFK LQLEEVQESPSKISINEINDQSDK DTELGENYSETSETYTELIEKTEKIKINPEINTG DMNEDKPSISGIKNPKQLNPTYR VSYDAYDRNKTLDKRLNKKWAPRQITEQYNFLD LDCVADINKTIQLWIGYISKQLID NKITITETPGYIERTLIGTVKLWLQNLSSSLDT LRSNKKLDGTTTTTVDILNKYEI AIRNEFSSMTTEVEEQNKEKITNRNMTKLAICN MCYIDEYTCAFRDYYYKGTYSPE SKEIRKLYFTKLPEPFSSKIISWNEAGLADTLG

VRIKFLQNWFIELCEKYKENMKME
 KILVKNLACCKSRIAPQFGCTDKYYKKEGKKKKF
 KSKYSKYK YRKP RRRYYVKNYKHK
 KPYRKKKKLTECTCYNCGKLGHLAKDCKLPKNPK
 KKQITEILIDNDKYTQVEYVDYEL
 SSED SIYEISENEFSENEINKDIEESDEENYD"
 CDS 2044..3321 /note="ORF2"
 /codon-start=1
 /product="putative movement
 protein"
 /protein-id="CAB42621.1"
 /db-xref="GI:4775495"
 /translation="MKKIMTEKDIQIIQQEEHQD
 EQSSEQKIIFDANIFEQIKGKELD
 LSIDKILEVPTIKNWFKRQKEEYVVSQREHIID
 CKYIKGKAQIPILNKRLLNKEIQD
 IKAKNPIKYVHLGGTEILIKACFREGIDTPIEY
 LADDRIVQPIEKSIISAVRGNLIY
 QKFKFIVSANYSVAVDDKNIDKSLVLYWKMSGIE
 LAPGSKIFTARCKNLYVLTTKHKI
 TAKNKINKIKIENPFERIVSVIDNNDYSYTEIDM
 DEDLEIVKERLSTSKRINNEMPET
 SSRSTSRSSTSKRINYTTTPQKLIEQKIEEINPHHY
 YITGIMDQRKYLILINTGQEENYV
 IRELIPEQEIVTIEQQNSELPKALRKNEETTEKE
 LIIGGIPILINFKIYQGDKNITLG
 IKWLEKVKPYKLED RQLTISYENKKIHKRTLI"
 CDS 3321..5231 /note="ORF3"
 /codon-start=1
 /product="putative polypeptide
 (aspartic proteinase, reverse
 transcriptase, ribonuclease H)"
 /protein-id="CAB42622.1"
 /db-xref="GI:4775496"
 /translation="MPKIYILSKIIVEGYYNRY
 TPMVDTGAEANMCRHNCLPESKWE
 KLKTPIVVTGFNNEGSMITYKARNIKIQIWDKIL
 TIEEISYEFTNKDILLGMPFLDK
 LYPHIITKTHWWFTTPCKQKLGA R VNNKVRKTT
 PWIKGSEKITQKLE NVIQSNHNIE
 IIFSINKIKPLQDKLELLYNDNPLQGWEKHQTK
 IKIELIDENSIITQKPLKYNFNDL
 TEFKMHIEKELLDNNYIQESNSKHTSPAFIVNKHS
 EQKRGKSRMVIDYRNLNAKTKTYN
 YPIPKNILKIRQIQGYNYFSKFDCKSGFYHLKLE
 DESKKLTAFTVPQGFYEWNVLPFG
 YKNAPGRYQHFMDNYFNQLENCIYIDDILLYSR
 TENEHIKLEKFIHIVEISGISLS
 KKKA EVMKNQIEFLGIQIDKNGIKMQTHVVQKII
 NLNETLDTKKKLQSFGLVNVQVRE
 YIPKLAENLKPLQKKLKKDIEYHFDEKDKIHIQK
 IKNMCKKLPKLYFPDEKKQFTYIV
 ETDSSDHSYGGVLKYKDNEKIEHHCRYYSGSYT
 EPQLKWEINRKELFGLYKCLLAFE
 PYIVYNKFIVRTDNTQVKWWITRKVQDSVTTKEI
 RRLVLNIQNFTFTIEVIRTDKNVI

CDS 5185..6393 ADYLSRQRYPNR"
 /note="ORF4"
 /codon-start=1
 /product="putative translation
 transactivator/ inclusion body
 protein"
 /protein-id="CAB42623.1"
 /db-xref="GI:4775497"
 /translation="MLLQTTYQDKGTQTDESQES
 KDIFTILTTLSQLMESMGKRLQQL
 ESQQHDYKNAELSQSEDSKLPEVEGDVGKLQKTH
 NTVALYTAAGTSKQVSKKPHINVN
 LNTVFDKPFTSKKPREAIVAPQTSTYANSLHHN
 KKVYNHITQTYIENIYKIQTFLNL
 NPRSTTTTDPDQDYVTQKLQGYNRLIAQPKTKAN
 LVKTCYNYGLLSTVYTHDGEEIIG
 IPELYKAFVTFKRITKGNLFFIKFYTAPAEILYD
 EIKPIQVIKIGLTRDMIPEEIE
 KQPEIQKMEIPSFYANKRIIGIATIIQELANNYL
 QGNAIWSYYSRDQLMIYANSKELR
 QGDMDEVQKWLSLLKPEMQPTTRALKKEFISNE
 LLTRYCKLVGHKYPDHICSKCNGD
 DNYVPEVQLE"
 repeat-region 6446..6661 /rpt-unit=6446..6521
 /rpt-type=TANDEM

SEQUENCE (SEQ):

1 tggatcaga gctatgacaa ataaacatat actttaagaa acaaaacat ggaaaggaac
 61 actgacatta atagtacaaa cccgcaaaac agtgaacaa tcaaaataac taagaaatcg
 121 cagaaaatta ggaagaaact aaagaaacta tatatagaat atgaatactt aagtatcaca
 181 aaaataaaca aagctaggct gtccaaatta atcgatataa tatctaaaac agaataataa
 241 tatatttggt atctaaagga taaaagatga ataaagaaga attcgatta gaagagaaaa
 301 catatgagaa tccagaagga ttaaaaataa caataatatt ttctaactta ggaagaagat
 361 ataaaaaat aggaataaac ctacttaaa tggtagaaaa agaactgta aaactagagg
 421 atagttaac cgccatggtt agaataacaa aagaaaacga agaatagat agaaaacgag
 481 agattaaaga aatacaacag caagctaaag aaaaaataca acagatagag gaagtaaaaa
 541 acactaaaat aacagaatta gaaaaagaat tagagatgct aaaacagatg tatgcaata
 601 aactaaaga aaaggaaaaa cgtaaagaag aagaagaaga actaaaacta acaaatgaga
 661 tagaaagatt caaattacag ttagaagaag tacaggagag tccatcaaaa ataagtataa
 721 acgaaataaa tgaccaaaagt gataagaca cagaactagg agaaaactat tcagaacaaa
 781 gtgaaacata cacagaactt atagaaaaaa cagaaaagat aaaaataaac ccagaataa
 841 atacaggaga tatgaacgaa gataaaccaa gcataatcagg aataaaaaat ccaaaacaat
 901 taaaccaac ctattacaga gtaagttagt atgcatatga cagaacaaa acattatggg
 961 ataaaagggt aaataagaaa tgggcaccaa gacagataac tgaacaatat aatttttag
 1021 atctagattg ttagcagat ataaataaaa caatacaatt atggatagga tatatctcaa
 1081 aacaactaat agataataaa ataacaataa cggaacacc aggatatata gaaagaacat
 1141 taataggaac tgtaaaatta tggttacaaa atttatcaag tgaagctta gacacattaa
 1201 ggagtaataa aaacttgac ggtacaacta caacaacagt tacagatata ttaataaat
 1261 atgaaatagc aataagaaat gaatttagta gtatgacaac agaagtagaa gaacaaaata
 1321 aagaaaaaat tacaataga aatttaataa caaaattagc aatatgtaat atgtgttata
 1381 tagatgaata tacttgtaga ttagagact attattataa aggaacatat agtccagatg
 1441 aaagtaaaga gataagaaaa ttatatttta caaaattacc agaacccttt agctcaaaaa
 1501 taataaaaag ttggaacgaa gcaggactag cagatacatt aggagtaagg ataaaattc
 1561 tacaactagc gtttatgaa ttatgtgaaa aatataaaga aaacatgaaa atggaaaaaa
 1621 tattagtaaa aaatttgga tgttgcaaaa gtagaatagc accccaattt ggtgcacag
 1681 ataaatatta caaaaaagaa ggaagaaaaa agaaatttaa atcaaaatat tcaagtata

1741 aatataggaa accaagaaga agatattatg taaaaaatta taacataaa aaacatata
1801 ggaaaaagaa aaaactaaca gaatgtactt gctataattg tggaaagcta ggacacttag
1861 ccaaagattg taaattacca aaaaacccaa agaagaaaca aattaccgaa atattgatag
1921 ataagataa atatacaca gtagaatatg tagattatga attagcagt gaagacagca
1981 tatatgatg atcagaaaac gagttctctg aaaatgaaat aaacaaggat atagaagagt
2041 cagatgaaga aaattatgac tgaaaaagat atacaaatta tacaacaaga agaaccacaa
2101 gatgaacaat cttcagaaca aaaaataata tttagcgcta atatatgtga acaataaaa
2161 ggaaagaat tggatctaag tatagacaaa atattagaag taccaacaat aaagaattgg
2221 tttaaagac aaaaagaaga atactatgta gtaagccaaa gagaacatat aatagattgt
2281 aaatacatca aaggtaaagc acaaatacca atttaataa aaagactact aaataaagaa
2341 atacaagata taaagcaaaa aaatccaata aaatatgtac acttaggagg aacggagatc
2401 ctaataaaag catgttttag ggaaggaata gataccccta tagaaatata ctggcagat
2461 gataggattg tacaacctat agaaaagagt ataataagt ctgtaagagg taacttaata
2521 taccaaaaat ttaaatttat agtaagtgt aactattcag tagcagtaga tgataaaaat
2581 atagataaat cattagtatt atactggaaa atgtctggaa tagaattagc accaggaagt
2641 aaaatattca cagcaagatg taaaaatcta tatgtcttaa caacaaaaca taagataaca
2701 gctaaaaata aaataataa aataaaaata gaaaatccat tcgaaagaat agtatcagtt
2761 atagacaaca atgattacag ttatacagaa atagacatgg atgaagattt agaaatagta
2821 aaagaagaat taagcacatc aaaacgaata aataatgaaa tgccagaaac atcatcaaga
2881 agtacatcaa gaagtacatc aaaaagaata aattatacca ctccacaaa attaatagaa
2941 caaaaaatag aagaataaaa cccacatcat tattatataa caggaataat ggaccaaaga
3001 aaatatttaa tactaataa tacagggcag gaagagaatt atgtataag agaactaata
3061 ccagaacaag agatagtaac aatagaacaa caaaattcag aactaccaa agcgtaaga
3121 aaaacgaag aaacaactga aaaaatata attattggag gaataccaat attaataat
3181 tttaaaatat atcaaggaga taaaaatatt acactaggga taaaatgggt agaaaaagtc
3241 aaacatata aattagaaga taggcaatta acaataagtt atgaaataa gaaaataata
3301 ataaaaagaa ctttgatata atgccaaga tatacatact ttccaaaata atagtagaag
3361 gatattataa tagatattat acacattatg tggatacagg agcagaagct aatatgtgta
3421 gacataattg ttaccagaa agtaaatggg aaaaagctaaa aacccccata gtagtaacag
3481 gatttaataa tgaaggaggt atgataacat ataaagcaag aaatataaaa atacaatat
3541 gggataaaa attaacata gaagaatat atagtacga attcacaat aaagatatat
3601 tattaggaat gccatttta gataaattat acccacaat tatacaaaa acacattggt
3661 ggtttactc cccgtgtaaa caaaaattag gagcaaaaag agtaataat aaagtaaga
3721 aaacaacacc ctggattaaa ggaagtgaag agattacca aaaattagaa aatgtaatac
3781 aaagtaacca taatatagag ataatcattt tctcaataa taagataaaa ccactacaag
3841 ataaactaga attactatat aatgataatc cactccaagg atgggaaaaa catcaacaa
3901 aaataagat tgaactaata gatgaaaata gcataatac aaaaaaacct ttaaaatac
3961 attttaatga ttaacagaa ttaaaatgc atataaaga attattgat aataactaca
4021 tacaagaag taatagtaaa catactagcc cagcatttat agtaataag catagtgaac
4081 aaaaagagg aaaaagccgt atggttatag attatagaaa cttaaatgca aaacaaaaa
4141 catataatta tccgatacca aataaaatc taaaattag acaatacaa gatatatac
4201 atttagtaa attgactgt aaatcaggat ttaccattt aaaactagaa gatgaatcta
4261 aaaagttaac agcattcaca gtaccacaag gattttacga atggaacgta ttacctttg
4321 gatataaaaa tgcaccagggt aggtatcaac attttatgga taattacttc aaacaattag
4381 aaaattgtat aatatatata gatgatata tgctatattc tagaacagag aacgaacata
4441 taaaactact agaaaaattc atacacattg tagaaatc aggaataagt ttaagtaaaa
4501 agaaagcaga agtaatgaaa aatcaaatag aatttttagg tatacaata gataaaaacg
4561 gaataaaaat gcaaacccat gtagtaca aaataattaa ctggaatgaa acacttgata
4621 caaaaaagaa gttacaatca ttttaggat tggtaacca agtaagagaa tatattccta
4681 aattagcaga aaacttaaaa ccattacaga aaaaattaaa aaaggacata gaatatcatt
4741 ttgacgaaaa agataaata catatacaga agataaaaaa tatgtgtaaa aaattacca
4801 aactataatt tccagatgaa aagaacaat ttacatatat ttagaaaact gattctagt
4861 atcacagcta cggaggaggt ctaaaatata aatatgataa tgaaaaaatt gaacaccatt
4921 gtagatatta tccggatca tacacggaac cacaataaa atgggaaata aataggaaag
4981 aactatttgg attatataa tgttttagg catttgagcc atatatgtt tataacaat
5041 ttatagtaag aacagataat acacaagtaa aatggtggat aactcggaat gtacaggatt

5101 cagtaactac aaaggaaata aggagacttg tattaatat acaaaatfff acatttaca
 5161 ttgaagtaac acgaactgac aagaatgta ttgcagacta cctatcaaga caaaggtagc
 5221 caaacagatg aaagccaaga atcaaaagac atatttaca tcttactac tctatcctta
 5281 cagatggaga gtatgggaaa aagattaca cagctagaaa gtcagcagca tgactataaa
 5341 aatgcggagc taagtcaatc ggaagactct aaactccag aggtagaagg agacgttggg
 5401 aaactccaaa aaaccataa cacagttgct ttatatacag ctgcaggtac aagcaacaa
 5461 gttagcaaaa agccacatat caatgtaaac ctataacccg tatttgataa gccatttaca
 5521 tcaaaaaagc caagagaagc aatagttata gcccacaaa cttaaccta tgctaatagc
 5581 ctacaccaca acaaaaaggt atataacct attactcaa catatattga gaatatata
 5641 aaaatccaga catttctgaa cctaaacct agatcaacta ctactacaga tccaacaca
 5701 gattatgtaa cccaaaaact acagggatat aataggctta tagccagcc aaaaacaaa
 5761 gcaaacctag taaaacatg ttacaactac ggactactta gcacagtata taccatgac
 5821 ggagaagaaa taattggaat accagagcta tacaagcat ttgtcacct caagagaatt
 5881 acaaaaggca acctattctt catcaaatc tatacagcac cagcggaaat actatagat
 5941 gaaataaaa ccataataa ggtgataaag ataggactaa cacgagatat gataataccg
 6001 gaagagatag agaacaacc ggagatacag aaaatggaga taccaagttt ctatgccaat
 6061 aaaagaataa ttggtatagc aactattat caagaactag ctaacaatta tctacaagga
 6121 aatgctatct ggagctacta ttccagagac cagttgatga tatatgcaa ctcaagga
 6181 ctacgacaag gagatatgga tgaagtcag aaatggattt tatcattat aaagccagaa
 6241 atgcaaccaa ctaccagagc attgaaaaa gaatttattt caaacgagtt attacaaga
 6301 tactgcaaac tagtaggaca caaatatcca gaccacatat gttcaaagt caacggagat
 6361 gataactatg taccagaagt ccaactagaa tgaaggtatc aacaagaaga caaagaaga
 6421 agacagctac tggaacata taatgtaaag taaatagtag tagtcacatt catgaacagt
 6481 aaaggtcgtt catgaatagt aagagtcagt taaaaagtc gtaagtaaa tagtatatgt
 6541 cataatcatg aacagtaaa gtcgttcagt aatagtaaga gtcagttaa aaagtcgtaa
 6601 agtaaatagt acgagtcata atcatgaaca gtaaggtcgt tcatgaata gtaggagta
 6661 ttgttaaaca gtaagagtcg tttaatttt ctttatatag aatgtaaat tgaggagga
 6721 ggacatcctc agacatcctc atcctcacct tctctctctt atctctctc aataaaat
 6781 ctgattatat acaaaacttt gaaagctatg gagcattcaa acgagttaca aaaccaggta
 6841 agttttatta taatcatgtt taactaaact ctatatcc ttataatc ttgaatatca
 6901 taattgttta tcatgttata gtactgttat aaagaacatc ttgagaaagt ttgctgtct
 6961 aataatgctg agagtagtta agcccagact atgccatcct aaagtggaaa ccagtaggca
 7021 aggaagccgt ttgggggagt acacagagtt gaggcgctgt tagggtaaat gattgaagca
 7081 tgaaaaacat ggtagtgtacc agaaaagatt gttcagtata actattaaa atagtataa
 7141 ctctatgata aacaaaggagg tttagaggaa tatgtttagt aaacacatga taaggataa
 7201 tagtaaatct gaataacaa ccacacgtat ttattagaag aaaatattga ctacaaaata
 7261 cttatgttat atactttta gagacttaac aacgatgtta aaagaaaaat ttacgatatt
 7321 ttagtacttt ttgaaataat atagttaag gaacaagcat ttacagaact aattgtatca
 7381 catgaatag atattataga tctatatgaa ctgatttt attcagacta atgacacat
 7441 atcaagttaa taaaatcttt ttataaatga aacataacac ttatcattt ataagattct
 7501 ataataccca gaaatggaaa cttcttaaaa acctcagtaa cataataga aaatagatt
 7561 gtgttaata caatattaaa acctgcaaag atattattag atataataga ttacgtaata
 7621 aagctttatt aactatagaa aatgaaattg aattttatga agataaactt gaaagttatc
 7681 aacacagaaa agaataact ctgagaataa tagaaacct ggatatatgt atcaacaggt
 7741 atacatctca gtattaatat gaacaaatca aaaattgata gttacaataa ttttaagtag
 7801 tggtagaaa ttctaaaaca tataaatatg tcaaaaataa tattagaaa aaatatagaa
 7861 gaatatcaaa aaactaaaga tattaatac ttagaatata cacttgaaaa tataaaagaa
 7921 attccaaca taaaaagat cctatgaaa ataagatgaa ctcaacaac ccccccaaa
 7981 t

L8 ANSWER 139 OF 146 GENBANK.RTM. COPYRIGHT 2001
 LOCUS (LOC): U95208 GenBank (R)
 GenBank ACC. NO. (GBN): U95208
 CAS REGISTRY NO. (RN): 197678-22-1
 SEQUENCE LENGTH (SQL): 7206
 MOLECULE TYPE (CI): DNA; linear

DIVISION CODE (CI): Viruses
 DATE (DATE): 28 Jun 2001
 DEFINITION (DEF): Petunia vein clearing virus, complete genome.
 SOURCE: petunia vein clearing virus.
 ORGANISM (ORGN): petunia vein clearing virus
 Viruses; Retrovird viruses; Caulimovirus
 NUCLEIC ACID COUNT (NA): 2383 a 1470 c 1285 g 2068 t
 COMMENT:

On Jun 28, 2001 this sequence version replaced gi:2463654.
 REFERENCE: 1 (bases 1 to 7206)
 AUTHOR (AU): Richert-Poggeler,K.R.; Shepherd,R.J.
 TITLE (TI): Petunia vein-clearing virus: a plant
 pararetrovirus with the core sequences for an
 integrase function
 JOURNAL (SO): Virology, 236 (1), 137-146 (1997)
 OTHER SOURCE (OS): CA 127:327178
 REFERENCE: 2 (bases 1 to 7206)
 AUTHOR (AU): Richert-Poeggeler,K.R.; Hohn,T.
 TITLE (TI): Isolation of an infectious full-length clone of petunia
 vein clearing virus (PVCV) from infected Nicotiana
 glutinosa
 JOURNAL (SO): Unpublished
 REFERENCE: 3 (bases 1 to 7206)
 AUTHOR (AU): Richert-Poeggeler,K.R.; Shepherd,R.J.
 TITLE (TI): Direct Submission
 JOURNAL (SO): Submitted (26-MAR-1997) Agronomy, University of
 Kentucky, Lexington, KY 40546-0091, USA
 REFERENCE: 4 (bases 1 to 7206)
 AUTHOR (AU): Richert-Poeggeler,K.R.; Shepherd,R.J.
 TITLE (TI): Direct Submission
 JOURNAL (SO): Submitted (28-JUN-2001) Agronomy, University of
 Kentucky, Lexington, KY 40546-0091, USA

FEATURES (FEAT):

Feature Key	Location	Qualifier
source	1..7206	/organism="petunia vein clearing virus" /specific-host="petunia hybrida cv. Himmelsroeschen" /db-xref="taxon:59504" /note="isolated from Nicotiana glutinosa which can be infected with the virus via grafting"
CDS	51..6590	/note="contains putative movement protein, RNA binding domain, protease, reverse transcriptase, and RNase H" /codon-start=1 /product="ORF I polyprotein" /protein-id="AAK68664.1" /db-xref="GI:14574598" /translation="MTSPSDYQSNSSLATTYSNA PKLSKALSNKYDYLYEVDILKENQ KISDTYLPLLPYSAFAKRSVTPWSQIRSLVQSK PRHVKEYVAASKLDQHPVFATGEE"

QFVTLHIPEEFASHWKSHQFTHIHFGAVKIALTY
HGRKGQPVVARLALLDTRYLEYQH
ANLGTAEITLNAGTVFITLFPNFTMSLSDANLST
ALKIQVQIQGAPLTKDSIQATLHY
QIAWRVQNHAMDLTLPGGEEALFLKIDAGSGATQ
CTQVPRQLSKEDLIKILPDSWVTN
YEKLREPEEPLRSTEVSMKRHDKSVAISFDHSH
YKKLRNTHHFMGMISDDVIVLDDP
ETFSKTLPSLMQTHDWIHHFQLDGRAVSWYKDPF
DGHCPWDIDCQCYSCLYSEDEEDF
EDGFPTKYKGIPRPGSIAERKMQEEANLKKLYEE
KDPFVGSLSRPGKYEYLVRDAPS
WAKDPHLTVEPTGWDSDEPIPPKQFPTTRNTLPK
IYMFNPLNYENNFPPLSSFSKDGA
DHTPKIPKRNVLPSGAKDPTGDLEATVNWQTEN
ALAQNRMLTTIDRTLKETVTKVDR
VTDQSSKNQGLIKVLEQQLQDLNKRICPPGTSLF
HFFDQQKSEMASLKEQIRLLKEQP
QKNETDTPSYQSSYQPFHSFSSPYMPSNPPNSPF
TTFANTPQPQPSLFSQYPIQPKSP
NTFDLAKLVWEKKDAIAAEKRAKKKLQKDEVKQK
TSLPESKRDPDQSSSHLGDQFMI
SDPALPKVYALNEPSVPS EDTSSQSYISTEESVE
DTDSFSVVSEESTQLSQLSSSSND
SPENNENTLPQTFMVRPTEPEISEVEDEVGDMTE
EPIPERRPEITPPKMVGTFHTFS
LDDISITKWPERIQDFHTWMLTKQLVEREPFLIL
SEFTARLSGTLREWWNSVGPDDKN
RFLTSQDFTWNIRILYSYFCGDQSQNKELRRQI
FEMKCLSYDRKKIDRHFQRMKLF
YHIGGDISLKQAFISSLPILSERISALIKERGT
SVTQMHVGDIRQTAFYVLDLCSK
RKFFNQMKKMSRDLEKACTKSDLIKGDKGCSGY
CNPSRRRKYKRFKLPSFKERDGRQ
YRKRFRFFRRSKTSKAMRQKPRSCFTCGKIGHFS
RNCPQNKKSIKLISEIQKYTGIDI
EDDLESVFSIEDEPSEDTLFSLEFYEEYAGEQYQ
ITSYEAPKTENPPLPKIHTIVEIP
QTEVKIYTSKWDPISVIAFYDTGAAYSIMDPAI
LPSEYWIPHFRHFGTADDGILTTT
VKT KHPITIEFFPGFKYTTKLLGSDIPGKDLLIG
FDIYRQLNNKLRIGADGIRWKNQF
KRYTEIPRLFQLTTSNELQQLEDVIKNQLCADSH
VDFLSKCSHPLWLNQDFFIQLPFK
KNENINPTKASHSGMNPEHLQLAIKECDELQQFD
LIEPSDSQWACEAFYVNRSEQVR
GKLRLVINYQPLNHFLQDDKFPIPN
KLTLFSLHSAKLFSKFDLKS GFVQLGIHPNER
PKTGFCIPDRHFQWKVMPFGLKTAPSLFQKAMIK
IFQPILFSALVYIDDILLFSETLE
DHIKLLNQFISLVKKFGVMLS AKKMILAQNKIQF
LGMDFA DGTFS PAGHISLELQKFP
DTNLSVKQIQQLGIVNYIRDFIPEVTEHISPLS
DMLKKKPPAWGKCQDNAV KQLKQL
AQQVKSLHIPSEGKKILQTDASDQYWSAVLLEEH
NGKRKICGFASGKFKVSEQHYHST

FKEILAVKNGIKKFNFLLIHTNFLVEMDMRAFPK
 MIRLNPKIVPNSQLLRWAQWFSPY
 QFEVKHLKGKDNILADFLSRPHEFSQRLKNSPKV
 LMFQRRTRSSSTKSKADSSQSTGS
 SYKLSHNLNPNPPEVFNLDYPWDTSVFLERRTFY
 ELQVFKKYGGSSILRPFQVDPEYFP
 AHIFIPNPTDFSEDLLWMFWYLLNHFHILMKFRC
 SKFSKIDQVNPWMLKFLWFWNNHN
 YWASLFCMKGIKKYVVIWFYRPVNYQQKLCAL
 PHSSIVKWNHVSVLNDEDEYSELQ
 RFIFQENKCIWPGSSGSWNYGNSDHPHGQW
 IRDALREYREMNDYFQDAQDPYPA
 YSKVDLTQEELNTRLRITRSYGSSSEDADMVKRSI
 YTVQSNIVKDSRKRKGKAKSRSS
 TRSEKRRAKNKCKYRSLHGEDWWIELGYSTKPST
 PSWTQDSSSEPCV"

SEQUENCE (SEQ):

1 tggatcaga gccatccatt ttgtgaggg aggaacatt tctaccata atgactctc
 61 ctctgatta tcaatccaat tctctctgg ctaccactta ttctaatga cctaaactct
 121 ctaaagcttt gtctaataaa tatgactacc ttatgaagt agatattcctt aaagaaaacc
 181 aaaaaatate agatattcctt ctctcttgg taaatcctta ctctgctttt gcaaagagat
 241 ccgttacccc ttggtctcag atccgttcat tagtccagtc taagcctcgc catgtgaaag
 301 aatatgttgc tgcctccaaa ttagatcagc atcctgtctt tgctacagga gaagaacagt
 361 ttgttactct acatattcca gaagaatttg ctctcactg gaaatccac caattcactc
 421 atatccactt tgggtctggt aaaattgccc ttactacca tggaagaaag ggtcaacctg
 481 ttgtgccag attagcactc ttagatacta gatatttga ataccaacat gctaatcttg
 541 gaactgctga aatcactctc aatgcaggca cagtcttcat tactctcttc cctaatctta
 601 ccatgtcttt atcagacgcc aatctgtcca cagccttaaa aatccaagtc caaatccaag
 661 gagctccctt tacaaaagac tccatccagg caactttaca ttaccaaatt gcttgccgtg
 721 tccaaaatca tgctatgat ttgactctcc ctggagggtga ggaagccttg ttctcaaaa
 781 ttgatccgg aagtggagcc acccaatgta cccaagtcct tagacagttg tcaaaagaag
 841 atctaataca aatattgcct gattcttggg taacaaatta tgaaggtta agggaaccag
 901 aagaacctct cgttccact gaagtctcaa tgtcaagag acatgataaa tcagtagcta
 961 tcagttttga tcaatccat tataagaagc tcagaacac ccatcacttc atgggtatga
 1021 ttcagatga tgcattgtc ttggatgatc ctgaaacttt ttccaaaact ctccatctc
 1081 tcatgcaaac tcatgattgg atccatcatt tccaacttga tggacgtgct gtatcatggt
 1141 ataaagatcc ttttgatggt cattgtccat gggatattga ttgccaatgt tattcttgtt
 1201 tatacagtg agatgaagag gatattgaag atggattcc taaaaatac aaagggattc
 1261 ccagaccagg aagtattgct gaaagaaga tgcaagaaga agctaattc aagaaacttt
 1321 atgaggagaa ggatcctttt gtgggatcat tgagcaggcc aggcataat gagtatctcg
 1381 ttagatacga tgcccatct tgggcaaaag atcctcatct gacagtggag ccaacagggt
 1441 gggatagtg tgaaccaatt ccactaagc aaccattcac aaccagaaat accttgccaa
 1501 aaatttacat gttcaacca ctcaattatg aaaacaactt tccaccttg tcatctttt
 1561 ctaaagatgg tgcagatcat actccaaaaa ttccaaagag aaatgttgtt ttaccatctg
 1621 gtgcaaaaga cccaactggt gatcttgagg caactgtaaa ttggcaaca gaaaatgctt
 1681 tggctcagaa taggatgtta acaaccattg ataggacct caaggagact gtaaccaagg
 1741 ttgatagagt cacagatcag tctcaaaaa atcaaggctt tatcaaggt ctgacaacac
 1801 agctcaaga ccttaacaag aggatgtgc cccaggaac atccttattt cacttcttg
 1861 atcaacaaaa gagtgaatg gcatccctta aagaacaaat cagactctc aaagaacaac
 1921 cacagaaaaa tgaactgat acccatctt accagtctc atatcagcct ttccacagt
 1981 tctctcccc atatagccc tcaatccgc ctaactcacc gttcacaact ttgcaata
 2041 tccacaacc acaactctt cttttctc aatatcaat acagccaaaa tcccaata
 2101 cctttgactt agccaaacta gtctgggaga aaaaagatgc aattgctga gaaaaagg
 2161 ccaagaagaa gctacaaaa gatgaggtta agcaaaaaac ttctttacct ccagaatcaa
 2221 aaaggccaga tcccagctt agcagtcacc ttggtatca gttcatgatt tctgatccag

2281 ctttaccaaa agtttacg cgatgaac cttcagttcc ttggaagat acatctccc
2341 aatcatacat cagtactgag gaatcagtag aagacacaga ctctttctct gtagtctcg
2401 aagaatcaac ccaactttct caactctcca gttcttctaa tgattacca gaaaacaatg
2461 aaaatacact acctcagacc ttatgggtc gtccaactga accggaaata tctgaagtg
2521 aagatgaggt agatggtatg actgaggaac caattctga aagaagacct gaaatcactc
2581 cacctaagat ggttggcaca ggcttcaca ctttcagcct agatgatatt tccattacca
2641 aatggcctga aagaattcag gatttcaca catggatgct cactaaaca ctggtgga
2701 gagaaccatt ctaatctcg tcagaattca cagcaagatt gtcaggaaca cttcgagaat
2761 ggtggaattc ttaggacca gatgacaaa acaggttctt gacttctcaa gactcacat
2821 ggaatattag aatactttat tcttattct gtggagatca aagtcaaac aaggaagaac
2881 tcaggagaca aatcttcgag atgaagtgtt tctctatga caggaagaaa atagatcgtc
2941 atttccaaag aatgatcaag ctcttttct acattggagg agatatttct taaagcaag
3001 ccttcatac ttccttccc cccattttat cagaaggat ctctgcctg ataaaagaaa
3061 ggggtacttc agttacaca atgcatgttg gagatattag gcaaacggct tttacgtcc
3121 ttgatgatc ctgtagcaaa agaaaattct tcaacaaat gaagaagatg tctagagatc
3181 tggaaaaagc atgcacaaa tctgacctta ttattaaggg tgataaagga tgttcaggct
3241 attgtaatcc ttcaagaaga aggaataca agcgcttcaa acttctagc ttcaaggaac
3301 gagatggcag gcaatacagg aaaagaagga ggttttcag aagaagcaaa acaagcaagg
3361 caatgaggca aaagcctaga tctgtttta catgtgggaa aattggtcac ttctccaga
3421 actgccccca aaacaagaag tcaataaaat tgatttctga aatccaaaag tatactggaa
3481 ttgatattga agtagacctg gagtccgtt tctcaattga ggatgaacc agtgaggata
3541 cattattctc gctagagttt tatgaagaat atgcaggaga gcaatccag atcacttct
3601 atgaagctcc aaaaacagaa aaccccccat tgcctaaaat ccacaccatt gttgaaatac
3661 cacaacaga ggtcaagatc tacacctcta aatgggataa gcctatctca gtcattgctt
3721 ttatgatac aggagcgcc taitcaatta tggatccagc aattcttcca tctgaatatt
3781 ggattccgca ttccaggcac ttggaacgg ctgatgatgg tatttgaca accacagtca
3841 aaaccaagca tccattaca attgaattct tcccaggctt caagtacact acaaaattgt
3901 tgggatcaga tttctctga aaagatctgc tcattggatt tgatatctac agacagctta
3961 ataataagct tcgattgga gcagatggaa tcagatggaa aaaccagttc aaaaggata
4021 ctgagatacc acgacttttc cagttaaca caagcaacga gcttcaacag ctgaggatg
4081 ttatcaaaaa tcaactctgt gcagattctc atgttgattt cctcagcaaa tgttctcctc
4141 cattatgct taatcaggat ttctcatcc agtccccct taaaaagaat gaaaacatca
4201 atccacaaa ggcagccat tctggaatga atccagaaca tctccagta gcaataaagg
4261 aatgtgatga gcttcaacaa ttgatctaa tgaaccctc tgattcaca tgggcatgtg
4321 aagcttttta tgtcaacaag aggtcagagc aagtgaagg aaaattaagg ttatcataa
4381 attatcaacc attaaacct tcttccaag atgacaagtt tccattcca aataagctca
4441 cactctctc tcatattcc aaagcaagc tctttcaaa gttgatctt aagtcaggat
4501 ttggcaatt aggaatccat cctaagaga gacaaaaac agggtttgc attcctgata
4561 gacatttca atggaaggc atgcatctg gtcttaaac agccccctt ttgttcaaa
4621 aagccatgat caaatcttc cagccatcc ttttctcagc ttggtctac attgatgata
4681 tcttctttt ttctgaaaca ctggaagatc acataaaat gcttaaccag ttcataagt
4741 tggtaaaaa attggagtc atgttatcg caaaaaagat gatccttct caaaaaaaa
4801 ttcaattctt aggaatggat ttgtctgat gaacttttc acctgtgtt cacatctct
4861 tggaaactca aaagtctct gatacaaat tctcggtaaa gcagataca caatttctg
4921 ggattgttaa ttatatcaga gattttatc cagaggttac tgagcatatt tcccctcta
4981 gtgatatgct aaagaagaaa ccaccggcat ggggaaaatg tcaagacaat gcagtcaagc
5041 aattgaagca actggctcaa caagtcaat cattacacat ccttctgaa ggaaagaaa
5101 tacttcaaac tgatccagat gatcagtatt ggagcgtgt actattggaa gaacataatg
5161 gcaaaaggaa aatctgtgga ttgcaagt gaaaatttaa ggtctctgaa caacattatc
5221 actccacatt caaagagatc ctgtctgta agaattgaat taaaaagttt aatttctct
5281 tgatccatac caacttttg gtcgaaatgg acatgagagc atttccaaag atgacaggat
5341 taaatctaa gattgttctt aattctcagt tgctacggtg ggcacaatgg ttttctcct
5401 accaatttga ggtaaaaat taaagggaagg agacaatat cctgtctgat ttctatcaa
5461 ggccacatga atttcacaa aggttaaga attcaccaa ggttctatg ttcaaagga
5521 ggacaagatc aagctcaaca aaaagcaagg ctgactcaag ccaaagtaca ggatcaagtt
5581 ataaattatc tcataatctt cctgaaaacc ctccggaagt tttaatacta gattatccat

5641 gggatacgtc tgttttcta gaaagaagaa cattctatga actacaagtt ttaagaaat
 5701 atggaggatc tattttaagg cctttggag ttgatccaga gtatccattt gtcacatct
 5761 ttatacctaa tcttacagat ttcttgaag atttattatg gatgttttg tatctgctca
 5821 atcatttcca catcctcatg aaattcaggt gtagtaaat cagcaagatt gatcaagtta
 5881 atccatggat gttgaagttc ttattatggt tcaataatca taattattgg gcttcattat
 5941 tcaagtgc atgagggaata aagaaatag tagtaattg gttttacaga ccagtaaatt
 6001 actaccaagg caagctttgt gctttacccc attcaagtat tgtgaagtgg aatcatgtct
 6061 ccgttttaaa tgatgaagat gaatactctg agctacaaag attcattttt caagagaaca
 6121 agtgattcc gaaagaaata tggccaggat cctcaggcag ttggaattat ggaaattcag
 6181 atcatcctca tggccaatgg attagagatg cacttaggga atacgagaa atgaacgatt
 6241 actttcaaga tgcacaagat ccatatcctg cttattctaa agttgatctc actcaagaag
 6301 aactaaacac ctacgtatc acgcgatcgt atggtagtag ttctgaagat gcagatatgg
 6361 tgaaaagaag tatttacaca gtccaatcca atatagtaaa ggattctcca agaaaaagaa
 6421 aagggaagc aaatcaaga tcatccacca gaagcgaaaa aagaagagct aaaaacaaat
 6481 gcaaatatcg gagtcttcat ggagaagatt ggtggattga attgggatat tctacaaagc
 6541 catcaacccc ctcattggaca caggacagct catcagaacc atgtgtctaa agcatgcgtg
 6601 aagaagataa gagggaatcc acttttcagt agtggctcta tgtgtctaaa gcatgcgtga
 6661 agaagataag agggaatcca cttttcagta gtggcttatt actttaata aagtttgtct
 6721 agtctacttc tgtctttatt aaaagtaagg aatctcttgc ttttctatt atggcatttt
 6781 gtaataatgc tctgtccact atcgagtctg tattattaag tgtgataagt gtgctgtcat
 6841 gatgtaagt cttacgtctt gagtcataga gattgcctat ataaggtaac ctcattagtt
 6901 gttttgatc agaaccaact ttctttgta aaatattatc atcaataaaa tgtttgtgtg
 6961 aataactctc tcttgtttcc aaactctcaa aatggtagtt cgaatcctaat attaaattga
 7021 agtgattac tctataagag tgattactga agatagtagg tcagattaag gttattgact
 7081 tggattatg gctgtatggt gccgtgttaa agcccagtag tcaacaaaa catctgatct
 7141 agttatctc agtttaatca gaattctctt ataggttcaa tcttataatt cctccatatc
 7201 tcgtat